

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 8, 2004, 11:00:44 ; Search time 28.7778 Seconds
(without alignments)
76.748 Million cell updates/sec

Title: US-09-753-139C-1

Perfect score: 42

Sequence: 1 CXCRPH 7

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_virus:*
16: sp_bacteriap:*
17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	40	95.2	189	4 Q9HD73	Q9hd73 homo sapien
2	40	95.2	445	4 Q96PY0	Q96py0 homo sapien
3	38	90.5	225	12 Q9IGT4	Q9igt4 porcine ade
4	38	90.5	527	5 P91773	P91773 penaeus jap
5	37	88.1	110	11 Q8BS77	Q8bs77 mus musculu
6	34	81.0	289	10 Q7XMI5	Q7xmi5 oryza sativ
7	33	78.6	190	11 Q8BJA8	Q8bja8 mus musculu
8	33	78.6	246	5 Q17792	Q17792 caenorhabdi
9	33	78.6	1376	12 Q83330	Q83330 murine hepa
10	33	78.6	1427	11 Q8YIB7	Q8yib7 mesocricetu
11	32	76.2	40	7 Q860N3	Q860n3 rattus sp.
12	32	76.2	50	9 Q854K7	Q854k7 mycobacteri
13	32	76.2	89	10 P82627	P82627 arabidopsis
14	32	76.2	101	5 Q8IMN0	Q8imn0 drosophila
15	32	76.2	103	17 Q9YGB0	Q9ygb0 aeropyrum p
16	32	76.2	119	5 Q76648	Q76648 caenorhabdi

17	32	76.2	134	4 Q8N788	Q8n788 homo sapien
18	32	76.2	193	4 Q8ETC4	Q8etc4 homo sapien
19	32	76.2	194	4 Q8BV29	Q8bv29 homo sapien
20	32	76.2	194	4 Q8NAR7	Q8nar7 homo sapien
21	32	76.2	246	10 Q8RYG1	Q8ryg1 oryza sativ
22	32	76.2	304	2 Q53405	Q53405 acetobacter
23	32	76.2	311	5 Q9W2C2	Q9w2c2 drosophila
24	32	76.2	335	12 Q8JXG9	Q8jxg9 heliothis z
25	32	76.2	337	4 Q8NHD3	Q8nhd3 homo sapien
26	32	76.2	342	4 Q8NHD5	Q8nhd5 homo sapien
27	32	76.2	344	2 Q84IK9	Q84ik9 streptomyce
28	32	76.2	351	4 Q8TF26	Q8tf26 homo sapien
29	32	76.2	353	16 Q8Y3G4	Q8y3g4 ralsstonia s
30	32	76.2	362	16 Q7V5J8	Q7v5j8 prochloroco
31	32	76.2	401	4 Q9H1U4	Q9h1u4 homo sapien
32	32	76.2	434	10 Q8S1U4	Q8s1u4 oryza sativ
33	32	76.2	442	10 Q8S1U7	Q8s1u7 oryza sativ
34	32	76.2	443	10 Q8S1U7	Q8s1u7 oryza sativ
35	32	76.2	461	5 Q9VZM1	Q9vzm1 drosophila
36	32	76.2	569	4 Q8NHD4	Q8nhd4 homo sapien
37	32	76.2	581	5 Q9BL07	Q9bl07 caenorhabdi
38	32	76.2	582	10 Q9AX44	Q9ax44 oryza sativ
39	32	76.2	600	11 Q8BH27	Q8bh27 mus musculu
40	32	76.2	610	16 Q83QL8	Q83ql8 shigella fl
41	32	76.2	647	5 Q86NWS	Q86nws drosophila
42	32	76.2	744	4 Q8NHD2	Q8nhd2 homo sapien
43	32	76.2	866	4 Q8IXF3	Q8ixf3 homo sapien
44	32	76.2	1275	11 Q9PFW0	Q9pfw0 rattus norv
45	32	76.2	1432	11 Q99J86	Q99j86 rattus norv

ALIGNMENTS

RESULT 1	Q9HD73	PRELIMINARY;	PRT;	189 AA.
AC	Q9HD73	01-MAR-2001 (TREMBLrel. 16, Created)		
DT	01-MAR-2001 (TREMBLrel. 16, Last sequence update)			
DT	01-OCT-2003 (TREMBLrel. 25, Last annotation update)			
DE	Omrithine decarboxylase antizyme.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.			
OX	NCBI_Taxid=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RA	Zhang W., Wan T., Cao X.;			
RL	Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.			
RL	EMBL; AF242521; AAF99601.1; -			
DR	GO; GO:0005509; F:calcium ion binding; IEA.			
DR	GO; GO:0008073; F:ornithine decarboxylase inhibitor activity; IEA.			
DR	InterPro; IPR002048; EF-hand.			
DR	InterPro; IPR002993; ODC_AZ.			
DR	Pfam; PF02100; ODC_AZ; 1.			
DR	ProDom; PD007483; ODC_AZ; 1.			
DR	PROSITE; PS00018; EF_HAND; 1.			
DR	PROSITE; PS01337; ODC_AZ; 1.			
SO	SEQUENCE 189 AA; 20999 MW; 232909D6AD01FEEF CRC64;			
QY	Query Match	95.2%;	Score 40;	DB 4; Length 189;
Db	Best Local Similarity	71.4%;	Pred. No. 9.8;	
	Matches 5; Conservative	0;	Mismatches 2;	Indels 0; Gaps 0;
QY	1 CXCRPH 7			
Db	31 CSCAPHP 37			
RESULT 2	Q96PY0	PRELIMINARY;	PRT;	445 AA.

AC Q96FY0;
 DT 01-DEC-2001 (TRENBLrel. 19, Created)
 DT 01-DEC-2001 (TRENBLrel. 19, Last sequence update)
 DT 01-OCT-2002 (TRENBLrel. 22, Last annotation update)
 DE Hypothetical protein KIAA1908 (Fragment).
 GN KIAA1908.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Brain;
 RX MEDLINE=21456161; PubMed=11572484;
 RA Nagase T., Kikuno R., Ohara O.;
 RT "Prediction of the coding sequences of unidentified human genes. XXI.
 RT The complete sequences of 60 new cDNA clones from brain which code for
 RT large proteins."
 RL DNA Res. 8:179-187(2001).
 DR EMBL; AB067495; BAB67801.1; -.
 KW Hypothetical protein.
 FT NON TER
 SQ SEQUENCE 445 AA; 47372 MW; FE0DF053B19AF891 CRC64;

Query Match 95.2%; Score 40; DB 4; Length 445;
 Best Local Similarity 71.4%; Pred. No. 19;
 Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CXCKP 7
 DB 48 CSCSP 54

RESULT 3

Q91GT4 PRELIMINARY; PRT; 225 AA.
 AC Q91GT4;
 DT 01-OCT-2000 (TRENBLrel. 15, Created)
 DT 01-OCT-2000 (TRENBLrel. 15, Last sequence update)
 DT 01-JUN-2003 (TRENBLrel. 24, Last annotation update)
 DE Iva2 (Fragment).
 GN Iva2.
 OS Porcine adenovirus type 3 (PAV-3).
 OC Viruses; dsDNA viruses, no RNA stage; Adenoviridae; Mastadenovirus.
 OX NCBI_TaxID=35265;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Agarwal N., Mittal S.K.;
 RT "Sequence Analysis of Porcine Adenovirus Type 3 E1 Region, PIX, pIVa2
 RT Genes, and Five Novel Open Reading Frames."
 RL Intervirology 0:0-0(2000).
 DR EMBL; AF247039; AAF78234.1; -.
 DR GO; GO:0019083; P:Viral transcription; IEA.
 DR InterPro; IPR003389; Adeno_Iva2.
 DR Pfam; PF02456; Adeno_Iva2; 1.
 FT NON TER
 SQ SEQUENCE 225 AA; 26243 MW; 112CDBB2E07F3B9 CRC64;

Query Match 90.5%; Score 38; DB 12; Length 225;
 Best Local Similarity 71.4%; Pred. No. 25;
 Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CXCKP 7
 DB 169 CTCPP 175

RESULT 4

P91773 PRELIMINARY; PRT; 527 AA.
 AC P91773;
 DT 01-MAY-1997 (TRENBLrel. 03, Created)
 DT 01-MAY-1997 (TRENBLrel. 03, Last sequence update)

DT 01-OCT-2003 (TRENBLrel. 25, Last annotation update)
 DE Pichi-2.
 OS Penaeus japonicus (Kuruma prawn).
 OC Eukaryota; Metazoa; Arthropoda; Crustacea; Malacostraca;
 OC Eumalacostraca; Eucarida; Decapoda; Dendrobranchiata; Penaeoidea;
 OC Penaeidae; Marsupenaeus.
 OX NCBI_TaxID=27405;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Watanabe T., Kono M.;
 RT "Isolation of a cDNA Encoding a Chitinase Family Protein from
 RT Cuticular Tissues of the Kuruma Prawn Penaeus japonicus."
 RL Zool. Sci. 0:0-0(1996).
 DR EMBL; D89751; BAA14014.1; -.
 DR GO; GO:0005576; C:extracellular; IEA.
 DR GO; GO:0005798; F:chitin binding; IEA.
 DR GO; GO:0005975; F:hydrolase activity, acting on glycosyl bonds; IEA.
 DR GO; GO:0006030; P:carbohydrate metabolism; IEA.
 DR InterPro; IPR002557; Chitin bind Pers.
 DR InterPro; IPR001223; Glyco_hydro_18.
 DR InterPro; IPR001579; Glyco_hydro_18AS.
 DR Pfam; PF01607; CBM_14; 1.
 DR Pfam; PF00704; Glyco_hydro_18; 1.
 DR ProDom; PD000471; Glyco_hydro_18; 1.
 DR SMART; SM00494; ChcBD2; 1.
 DR SMART; SM00636; Glyco_18; 1.
 DR PROSITE; PS01095; CHITINASE_18; 1.
 KW Glycosidase; Hydrolase.
 SQ SEQUENCE 527 AA; 59162 MW; B9CBABAB8CDF8710 CRC64;

Query Match 90.5%; Score 38; DB 5; Length 527;
 Best Local Similarity 71.4%; Pred. No. 49;
 Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CXCKP 7
 DB 500 CGCEP 506

RESULT 5

Q8BS77 PRELIMINARY; PRT; 110 AA.
 AC Q8BS77;
 DT 01-MAR-2003 (TRENBLrel. 23, Created)
 DT 01-MAR-2003 (TRENBLrel. 23, Last sequence update)
 DT 01-MAR-2003 (TRENBLrel. 23, Last annotation update)
 DE Mus musculus (Mouse).
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sclurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Embryo;
 RX MEDLINE=22354683; PubMed=12466851;
 RA The FANTOM Consortium.
 RA the RIKEN Genome Exploration Research Group Phase I & II Team;
 RT "Analysis of the mouse transcriptome based on functional annotation of
 RT 60,770 full-length cDNAs."
 RL Nature 420:563-573(2002).
 DR EMBL; AK035013; BAC28910.1; -.
 KW Hypothetical protein
 SQ SEQUENCE 110 AA; 12073 MW; B956CD7C1FC894C CRC64;

Query Match 88.1%; Score 37; DB 11; Length 110;
 Best Local Similarity 71.4%; Pred. No. 21;
 Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CXCKP 7
 DB 79 CWCOP 85

RESULT 6

07XMI5 PRELIMINARY; PRT; 289 AA.
 AC 07XMI5;
 DT 01-OCT-2003 (TREMBlrel. 25, Created)
 DT 01-OCT-2003 (TREMBlrel. 25, Last sequence update)
 DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
 DE OSUNBA0084K01.16 protein.
 GN OSUNBA0084K01.16
 OS Oryza sativa (Rice).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 OC Ehrhacridae; Oryzaceae; Oryza.
 NCBI_TaxID=4530;
 RN (1)
 RP SEQUENCE FROM N.A.
 RA Han B., Feng Q., Huang Y.C., Li Y., Zhu J.J., Zhao Q., Hu X.,
 RA Liu Y.L., Mu Z., Xu Z., Chen L., Fan D.L., Weng Q.J., Zhang L.,
 RA Lu Y.Q., Yu S.L., Liu X.H., Lu T.T., Zhang Y.J., Lu Y., Li C., Li T.,
 RA Zhang Y., Hu H., Jia P.X., Qian Y.M., Ying K., Zhou B., Chen Z.H.,
 RA Hao P., Zhang L., Wu M., Zhang R.Q., Guan J.P., Fu G., Wang S.Y.,
 RA Ren S.X., Lv G., Lin W., Gu W.Q., Zhu G.F., Tu Y.F., Jia J., Yin H.F.,
 RA Zhang Y., Cai Z., Chen J., Kang H., Chen X.Y., Shao C.Y., Sun Y.,
 RA Hu Q.P., Zhang X.L., Zhang W., Wang L.J., Ding C.W., Sheng H.H.,
 RA Gu J.L., Chen S.T., Ni L., Zhu P.H., Hong G.F.,
 RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AL606999; CAE04844.1; -
 SQ SEQUENCE 289 AA; 31003 MW; 2E52E631ACB445C3 CRC64;

Query Match Best Local Similarity 81.0%; Score 34; DB 10; Length 289;
 Matches 4; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 CXCPHP 7
 Db 279 CSCPYP 285

RESULT 7

08BJA8 PRELIMINARY; PRT; 190 AA.
 AC 08BJA8;
 DT 01-MAR-2003 (TREMBlrel. 23, Created)
 DT 01-MAR-2003 (TREMBlrel. 23, Last sequence update)
 DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)
 DE Hypochemical ankyrin repeat region circular profile/yeast DNA-binding
 OS Mus musculus (Mouse)
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 NCBI_TaxID=10090;
 RN (1)
 RP SEQUENCE FROM N.A.
 RA STRAIN=NOD; TISSUE=Spleen;
 RC MEDLINE=22354683; PubMed=12466851;
 RA The FANTOM Consortium;
 RA the RIKEN Genome Exploration Research Group Phase I & II Team;
 RT "Analysis of the mouse transcriptome based on functional annotation of
 60,770 full-length cDNAs.";
 RT Nature 420:563-573 (2002)
 RL EMBL; AK089667; BAC40941.1; -
 DR InterPro; IPR002110; ANK.
 DR PROSITE; PS0297; ANK_REGION; 1.
 DR Hypochemical protein.
 SQ SEQUENCE 190 AA; 19996 MW; 2D11E31D709B38FA CRC64;

Query Match Best Local Similarity 78.6%; Score 33; DB 11; Length 190;
 Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CXCPHP 6

Db 163 CACAPH 168

RESULT 8

017792 PRELIMINARY; PRT; 246 AA.
 AC 017792;
 DT 01-NOV-1996 (TREMBlrel. 01, Created)
 DT 01-MAR-2003 (TREMBlrel. 23, Last sequence update)
 DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
 DE Hypochemical protein C07E3.6.
 GN C07E3.6.
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
 OC Rhabditidae; Felodectinae; Caenorhabditis.
 NCBI_TaxID=6239;
 RN (1)
 RP SEQUENCE FROM N.A.
 RA Matthews P.;
 RL Submitted (JUN-1995) to the EMBL/GenBank/DBJ databases.
 RN (2)
 RP SEQUENCE FROM N.A.
 RA MEDLINE=99069613; PubMed=9851916;
 RX none;
 RT "Genome sequence of the nematode C.elegans: A platform for
 investigating biology."
 RL Science 282:2012-2018 (1998).
 DR EMBL; Z49908; CAA90100.2; -
 DR PIR; T19056; T19056.
 DR WormBep; C07E3.6; CE32306.
 KW Hypochemical protein.
 SQ SEQUENCE 246 AA; 26490 MW; D7EB56BA33F534A9 CRC64;

Query Match Best Local Similarity 78.6%; Score 33; DB 5; Length 246;
 Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CXCPHP 6
 Db 49 CSCPYP 54

RESULT 9

083330 PRELIMINARY; PRT; 1376 AA.
 AC 083330;
 DT 01-NOV-1996 (TREMBlrel. 01, Created)
 DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
 DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
 DE S glycoprotein.
 OS Murine hepatitis virus.
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Nidovirales;
 OC Coronaviridae; Coronavirus.
 NCBI_TaxID=11138;
 RN (1)
 RP SEQUENCE FROM N.A.
 RA STRAIN=RI;
 RC MEDLINE=95304829; PubMed=7785316;
 RA Kunlita S., Zhang L., Homberger F.R., Compton S.R.;
 RT "Molecular characterization of the S proteins of two enterotropic
 murine coronavirus strains";
 RL Virus Res. 35:277-289 (1995).
 DR EMBL; U14645; AA87062.1; -
 DR InterPro; IPR002552; Corona_S2.
 DR Pfam; PF01601; Corona_S2; 1.
 SQ SEQUENCE 1376 AA; 151615 MW; E9EE02151BBD7D9D CRC64;

Query Match Best Local Similarity 78.6%; Score 33; DB 12; Length 1376;
 Matches 4; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 CXCPHP 7

Db 532 CTCAPNP 538

RESULT 10

08VIB7 PRELIMINARY; PRT; 1427 AA.
AC 08VIB7;
DT 01-MAR-2002 (TREMBLrel. 20, Created)
DT 01-MAR-2002 (TREMBLrel. 20, Last sequence update)
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DE Attractin.
GN ATRN.
OS Mesocricetus auratus (Golden hamster).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;
OC Mesocricetus.
NC NCB1_Taxid=10036;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21635551; PubMed=1173967;
RA Kuramoto T., Nomoto T., Fujiwara A., Mizutani M., Sugimura T.,
RA Ushijima T.;
RT "insertional mutation of the Attractin gene in the black tremor hamster.";
RL Mamm. Genome 13:36-40(2002).
CC -1- SIMILARITY: CONTAINS 1 CUB DOMAIN.
DR EMBL; AB062913; BAB72012.1; -.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0004872; F:receptor activity; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR GO; GO:0005529; F:sugar binding; IEA.
DR GO; GO:0007275; F:development; IEA.
DR InterPro; IPR000859; CUB.
DR InterPro; IPR006529; EGF-like.
DR InterPro; IPR006552; Kelch_rep.
DR InterPro; IPR02049; Laminin_EGF.
DR InterPro; IPR001304; Lectin_C.
DR InterPro; IPR003659; Plexin-like.
DR InterPro; IPR002165; Plexin_repeat.
DR Pfam; PF00431; CUB; 1.
DR Pfam; PF01344; Kelch; 6.
DR Pfam; PF00053; Laminin_EGF; 1.
DR Pfam; PF00059; Lectin_C; 1.
DR Pfam; PF01437; PSI; 4.
DR PRINTS; PRO0011; EGF/LAMININ.
DR SMART; SM00034; CLECT; 1.
DR SMART; SM00042; CUB; 1.
DR SMART; SM00423; PSI; 5.
DR PROSITE; PS01180; CUB; 1.
DR PROSITE; PS00041; C_TYPE_LECTIN_2; 1.
DR PROSITE; PS00022; EGF_1; 3.
DR PROSITE; PS01186; EGF_2; 1.
DR PROSITE; PS01248; LAMININ_TYPE_EGF; 1.
KW Laminin EGF-like domain.
SQ SEQUENCE 1427 AA; 158024 MW; 31F865993BA17B68 CRC64;

Query Match 78.6%; Score 33; DB 11; Length 1427;
Best Local Similarity 66.7%; Pred. No. 7.7e+02;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CXCPH 6
Db 305 CSCSPH 310

RESULT 11
Q860N3 PRELIMINARY; PRT; 40 AA.
AC Q860N3;
DT 01-JUN-2003 (TREMBLrel. 24, Created)
DT 01-JUN-2003 (TREMBLrel. 24, Last sequence update)
DT 01-JUN-2003 (TREMBLrel. 24, Last annotation update)
DE Major histocompatibility complex class I Eu-like protein

DE (Fragment).

OS Rattus sp.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCB1_Taxid=10118;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96217551; PubMed=6613150;
RA Yuan X.J., Salgar S.K., Hassett A.L., McHugh K.P., Kunz H.W.,
RA Gill T.J., Iit.;
RT "Physical mapping of the E/C and grc regions of the rat major histocompatibility complex.";
RL Immunogenetics 44:9-18(1996).
DR EMBL; S81844; AAP32239.1; -.
FT NON_TER 1 40 1
SQ NON_TER 40 40
SQ SEQUENCE 40 AA; 4957 MW; FF31A75A18D217E3 CRC64;

Query Match 76.2%; Score 32; DB 7; Length 40;
Best Local Similarity 66.7%; Pred. No. 65;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CXCPH 6
Db 10 CTCGPH 15

RESULT 12

0854K7 PRELIMINARY; PRT; 50 AA.
AC 0854K7;
DT 01-JUN-2003 (TREMBLrel. 24, Created)
DT 01-JUN-2003 (TREMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DE Gp37.
OS Mycobacteriophage Omega.
OC Viruses; dsDNA viruses, no RNA stage; Caudovirales; Siphoviridae.
OX NCB1_Taxid=205879;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22592660; PubMed=12705866;
RA Pedulla M.L., Ford M.E., Houtz J.M., Karthikeyan T., Wadsworth C.,
RA Lewis J.A., Jacobs-Sera D., Falbo J., Grosse J., Pannunzio N.R.,
RA Brucker W., Kumar V., Kandamany J., Keenan L., Bardarov S.,
RA Kravkov J., Lawrence J.G., Jacobs W.R. Jr., Hendrix R.W.,
RA Hatfull G.F.;
RT "Origins of highly mosaic mycobacteriophage genomes.";
RL Cell 113:171-182(2003).
DR EMBL; AY129338; AAN12701.1; -.
DR GO; GO:0004872; F:receptor activity; IEA.
DR InterPro; IPR001368; TNFR_c6.
DR PROSITE; PS00652; TNFR_NGFR_1; 1.
SQ SEQUENCE 50 AA; 5273 MW; F5D9305CDF1EF00 CRC64;

Query Match 76.2%; Score 32; DB 9; Length 50;
Best Local Similarity 66.7%; Pred. No. 78;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CXCPH 6
Db 13 CTCGPH 18

RESULT 13

P82627 PRELIMINARY; PRT; 89 AA.
AC P82627;
DT 01-JUN-2002 (TREMBLrel. 21, Created)
DT 01-JUN-2002 (TREMBLrel. 21, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Hypothetical protein SCRL8 precursor.
GN SCRL8.
OS Arabidopsis thaliana (Mouse-ear cress).

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
 OC euroside II; Brassicales; Brassicaceae; Arabidopsis;
 NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. Columbia;
 RX MEDLINE=21016719; PubMed=1130712;
 RA Theologis A., Becker J.R., Palm C.J., Federspiel N.A., Kaul S.,
 White O., Alonso J., Altieri H., Arujo R., Bowman C.L., Brooks S.Y.,
 Buehler E., Chan A., Chao Q., Chen H., Cheuk R.F., Chin C.W.,
 Chung M.K., Conn L., Conway A.B., Conway A.R., Creasy T.H., Dewar K.,
 Dunn P., Egu P., Feldblum T.V., Feng J.-D., Fong B., Fujii C.Y.,
 Gill J.R., Goldsmith A.D., Haas B., Hansen N.F., Hughes B., Huzar L.,
 Hunter J.L., Jenkins J., Johnson-Hopson C., Khan S., Khaykin E.,
 Kim C.J., Koo H.L., Kremenetskaia I., Kuritz D.B., Kuan A., Lam B.,
 Langin-Hooper S., Lee A., Lee J.M., Lenz C.A., Li J.H., Li Y.-P.,
 Lin X., Liu S.X., Liu Z.A., Luros J.S., Maiti R., Marziani A.,
 Miltischer J., Miranda M., Nguyen M., Nieman W.C., Osborne B.I.,
 Pal G., Peterson J., Pham P.K., Rizzo M., Rooney T., Rowley D.,
 Sakano H., Salberg S.L., Schwartz J.R., Shin P., Southwick A.M.,
 Sun H., Tallon L.O., Tamunga G., Tortum M.O., Town C.D.,
 Utterback T., Van Aken S., Vaysberg M., Vysotskaia V.S., Walker M.,
 Wu D., Yu G., Fraser C.M., Venter J.C., Davis R.W.;
 RT "Sequence and analysis of chromosome 1 of the plant Arabidopsis
 thaliana."
 RL Nature 408:816-820(2000).
 RN [2]
 RP IDENTIFICATION.
 RX PubMed=11437247;
 RA Vanosthuisen V., Mlege C., Dumae C., Cock J.M.;
 RT "Two large Arabidopsis thaliana gene families are homologous to the
 Brassica gene superfamily that encodes pollen coat proteins and the
 male component of the self-incompatibility response."
 DR Plant Mol. Biol. 46:17-34(2001).
 KW EMBL; AC018908; -; NOT ANNOTATED CDS.
 FT SIGNAL
 FT CHAIN 1 23
 FT SIGNAL
 FT CHAIN 24 89
 FT SIGNAL
 FT CHAIN 89 AA; EBI1B65AEM497F CRC64;
 SQ SEQUENCE
 Query Match 76.2%; Score 32; DB 10; Length 89;
 Best Local Similarity 57.1%; Pred. No. 1.2e+02;
 Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
 QY 1 CXCXHP 7
 DB 67 CTCRPP 73
 RESULT 14
 Q8IMNO PRELIMINARY; PRT; 101 AA.
 AC Q8IMNO;
 DT 01-MAR-2003 (TrEMBLrel. 23, Created)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
 DE CG31308-PA.
 GN CG31308.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 NCBI_TaxID=7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=20196006; PubMed=10731132;
 RA Adams M.D., Celinker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 Amanatides P.G., Scher S.E., Li P.W., Hoskins R.A., Galle R.F.,
 George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 Sutton G.G., Mortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers Y.H., Blazek R.G., Champe M., Pfeiffer B.D.,
 Man K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Gabor G.L.,

RA Abril J.F., Agbayani A., An H.J., Andrews-Pfannkoch C., Baldwin D.,
 Ballew R.M., Basu P., Bendale J., Bayraktaroglu L., Beasley E.M.,
 Beeson K.Y., Bence P.V., Berman B.P., Bhandari D., Bolchakov S.,
 Borokova D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,
 RA Burris K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
 RA Foster C., Gabrielian A.B., Garg N.S., Gelbart W.M., Glasser K.,
 RA Glodok A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howland T.J., Wei M.H., Idegawa C.,
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kensington J.A., Kelchum K.A.,
 RA Kimmel B.E., Kodali C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Lasako P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merkulov G., Mishina N.V., Mobarry C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nuskern D.R., Pacle J.M.,
 RA Palazolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spier E., Spreading A.C., Stapleton M., Strong R., Sun E.,
 RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.Y., Wasserman D.A., Weinstein G.M., Weissenbach J.,
 RA Williams S.M., Woodger, Worley K.C., Wu D., Yang S., Yao Q.A., Ye J.,
 RA Yeh R.F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RT "The genome sequence of Drosophila melanogaster."
 RL Science 287:2185-2195(2000).
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Celinker S.E., Adams M.D., Krommiller B., Man K.H., Holt R.A.,
 RA Evans C.A., Gocayne J.D., Amanatides P.G., Brandon R.C., Rogers Y.,
 RA Banton J., An H., Baldwin D., Banton J., Beeson K.Y., Busam D.A.,
 RA Carlson J.W., Center A., Champe M., Davenport L.B., Dietz S.M.,
 RA Dodson K., Dorsett V., Doup L.E., Doyle C., Dreenek D., Farfan D.,
 RA Ferreira S., Frise E., Galle R.F., Garg N.S., George R.A.,
 RA Gonzalez M., Houck J., Hoskins R.A., Hostin D., Howland T.J.,
 RA Ibegwam C., Jalali M., Kruse D., Li P., Mattei B., Moshrefi A.,
 RA McIntosh T.C., Moy M., Murphy B., Nelson K.A., Nunoo J.,
 RA Pacle J.M., Paragas V., Park S., Patel S., Pfeiffer B., Scheeler F.,
 RA Phouenavong S., Pittman G.S., Puri V., Richards S., Scheeler F.,
 RA Stapleton M., Strong R., Svirskas R., Tector C., Tyler D.,
 RA Williams S.M., Zaveri J.S., Smith H.O., Venter J.C., Rubin G.M.;
 RT "Sequencing of Drosophila melanogaster genome."
 RL Submitted (MAR-2000) to the EMBL/Genbank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RA Mira S., Crosby M.A., Matthews B.B., Bayraktaroglu L., Campbell K.,
 RA Hradecky P., Huang Y., Kaminker J.S., Prochuk S.E., Smith C.D.,
 RA Tupy J.L., Bergman C., Berman B., Carlson J.W., Celinker S.E.,
 RA Clamp M., Drysdale R., Emmert D., Frise E., de Grey A., Harris N.,
 RA Krommiller B., Marshall B., Milburn G., Richter U., Russo S.,
 RA Seattle S.M.J., Smith E., Shu S., Smutnick F., Whitfield E.,
 RA Ashburner M., Gelbart W.M., Rubin G.M., Mungall C.J., Lewis S.E.;
 RT "Annotation of Drosophila melanogaster genome."
 RL Submitted (MAR-2000) to the EMBL/Genbank/DBJ databases.
 RN [4]
 RP SEQUENCE FROM N.A.
 RA Adams M.D., Celinker S.E., Gibbs R.A., Rubin G.M., Venter J.C.;
 RL Submitted (MAR-2000) to the EMBL/Genbank/DBJ databases.
 RN [5]
 RP SEQUENCE FROM N.A.
 RA Flybase;
 RL Submitted (SEP-2002) to the EMBL/Genbank/DBJ databases.
 DR EMBL; A0003762; A001430.1;
 DR Flybase; FBgn0051308; CG31308.
 SQ SEQUENCE 101 AA; 11247 MW; 766D30410CFA209 CRC64;
 Query Match 76.2%; Score 32; DB 5; Length 101;

Best Local Similarity 66.7%; Pred. No. 1.4e+02;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Oy 1 CXCPH 6
| | |
Db 77 CVCAPH 82

RESULT 15

O9YG80 PRELIMINARY; PRT; 103 AA.
AC O9YG80; 01-NOV-1999 (TRENBLrel. 12, Created)
DT 01-NOV-1999 (TRENBLrel. 12, Last sequence update)
DT 01-JUN-2003 (TRENBLrel. 24, Last annotation update)
DE Hypothetical protein APE0021.
GN APE0021.
OS Aeropyrum pernix.
OC Archaea; Crenarchaeota; Thermoprotei; Desulfurococcales;
OC Desulfurococcales; Aeropyrum.
OX NCBI_TaxID=56636;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-K1;
RX MEDLINE=99310339; PubMed=10382966;
RA Kawarabayashi Y., Hino Y., Horikawa H., Yamazaki S., Haikawa Y.,
RA Jin-no K., Takahashi M., Sekine M., Baba S.-I., Ankaï A., Kosugi H.,
RA Hosoyama A., Fukui S., Nagai Y., Nishijima K., Nakazawa H.,
RA Takamiya M., Masuda S., Funahashi T., Tanaka T., Kudoh Y.,
RA Yamazaki J., Kushiida N., Oguchi A., Aoki K.-I., Kubota K.,
RA Nakamura Y., Nomura N., Sako Y., Kikuchi H.;
RT "Complete genome sequence of an aerobic hyper-thermophilic
RT crenarchaeon, Aeropyrum pernix K1.";
RL DNA Res. 6:83-101(1999).
DR EMBL: AP000058; BAA78930.1; -.
DR PIR: H72753; H72753.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 103 AA; 10953 MW; 84D9605698B6A996 CRC64;

Query Match 76.2%; Score 32; DB 17; Length 103;

Best Local Similarity 57.1%; Pred. NO. 1.4e+02;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Oy 1 CXCPH 7
| | |
Db 13 CSCCTHP 19

Search completed: April 8, 2004, 11:16:13
Job time : 29.7778 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 8, 2004, 11:16:30 ; Search time 30.333 Seconds
(without alignments)

60.678 Million cell updates/sec

Title: US-09-753-139c-1

Perfect score: 42

Sequence: 1 CKCXPHP 7

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1073127 seqs, 262937947 residues

Total number of hits satisfying chosen parameters: 1073127

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications AA.*

```
1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep.*
10: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep.*
11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep.*
12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
13: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pep.*
14: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep.*
15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep.*
16: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
17: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
18: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	39	92.9	30	14	US-10-011-585A-157
2	39	92.9	89	12	US-10-424-599-254963
3	39	92.9	519	15	US-10-108-260A-767
4	38	90.5	170	10	US-09-753-139c-1
5	37	88.1	170	15	US-10-264-237-1473
6	37	81.0	289	12	US-10-412-699B-1681
7	34	81.0	289	15	US-10-374-780A-1670
8	34	78.6	91	12	US-10-424-599-153612
9	33	78.6	97	9	US-09-925-297-595
10	33	78.6	108	9	US-09-796-692-1071
11	33	78.6	108	14	US-10-040-862-1071
12	33	78.6	108	15	US-10-057-475B-1071
13	33	78.6	108	15	US-10-154-884B-1071
14	33	78.6	119	9	US-09-796-692-1736
15	33	78.6	119	14	US-10-040-862-1736

16	33	78.6	119	15	US-10-057-475B-1736	Sequence 1736, Ap
17	33	78.6	119	15	US-10-154-884B-1736	Sequence 1736, Ap
18	33	78.6	120	9	US-09-796-692-1065	Sequence 1065, Ap
19	33	78.6	120	14	US-10-040-862-1065	Sequence 1065, Ap
20	33	78.6	120	15	US-10-057-475B-1065	Sequence 1065, Ap
21	33	78.6	120	15	US-10-154-884B-1065	Sequence 1065, Ap
22	33	78.6	139	9	US-09-796-692-755	Sequence 755, App
23	33	78.6	139	14	US-10-040-862-755	Sequence 755, App
24	33	78.6	139	15	US-10-057-475B-755	Sequence 755, App
25	33	78.6	139	15	US-10-154-884B-755	Sequence 755, App
26	33	78.6	149	9	US-09-796-692-1077	Sequence 1077, Ap
27	33	78.6	149	14	US-10-040-862-1077	Sequence 1077, Ap
28	33	78.6	149	15	US-10-057-475B-1077	Sequence 1077, Ap
29	33	78.6	149	15	US-10-154-884B-1077	Sequence 1077, Ap
30	33	78.6	155	9	US-09-796-692-1642	Sequence 1642, Ap
31	33	78.6	155	14	US-10-040-862-1642	Sequence 1642, Ap
32	33	78.6	155	15	US-10-057-475B-1642	Sequence 1642, Ap
33	33	78.6	155	15	US-10-154-884B-1642	Sequence 1642, Ap
34	33	78.6	592	14	US-10-288-556-18	Sequence 18, Appl
35	33	78.6	665	12	US-10-424-599-149955	Sequence 149955,
36	32	76.2	40	10	US-09-820-649-143	Sequence 143, App
37	32	76.2	40	14	US-10-160-162-143	Sequence 143, App
38	32	76.2	57	12	US-10-424-599-262948	Sequence 262948,
39	32	76.2	114	12	US-10-424-599-223258	Sequence 223258,
40	32	76.2	133	12	US-10-001-843-158	Sequence 158, App
41	32	76.2	135	13	US-09-925-297-538	Sequence 538, App
42	32	76.2	149	9	US-10-106-698-5783	Sequence 5783, Ap
43	32	76.2	164	14	US-10-029-386-33998	Sequence 33998, A
44	32	76.2	172	14	US-10-029-386-34054	Sequence 34054, A
45	32	76.2	173	14	US-10-029-386-34054	Sequence 34054, A

ALIGNMENTS

```
RESULT 1
US-10-011-585A-157
; Sequence 157, Application US/10011585A
; Publication No. US2003003986A1
; GENERAL INFORMATION:
; APPLICANT: Sun, Yongming
; APPLICANT: Recipon, Herve
; APPLICANT: Chen, Sei-Yu
; APPLICANT: Liu, Chenghua
; TITLE OF INVENTION: Compositions and Methods Relating to Prostate Specific
; FILE REFERENCE: DEX-0261
; CURRENT APPLICATION NUMBER: US/10/011,585A
; CURRENT FILING DATE: 2002-03-14
; PRIOR APPLICATION NUMBER: 60/245,740
; PRIOR FILING DATE: 2000-11-03
; NUMBER OF SEQ ID NOS: 245
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 157
; LENGTH: 30
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-011-585A-157

Query Match          92.9% Score 39; DB 14; Length 30;
Best Local Similarity 71.4% Pred. No. 25;
Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY      1 CKCXPHP 7
      |||||
      24 CACFPHP 30

RESULT 2
US-10-424-599-254963
; Sequence 254963, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
```

```
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovacic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285584
; SEQ ID NO 254963
; LENGTH: 89
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_72254C.1.pap
US-10-424-599-254963

Query Match
Best Local Similarity 92.9%; Score 39; DB 12; Length 89;
Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 CXCRPH 7
Db 53 CTCRPH 59

RESULT 3
US-10-108-260A-3767
; Sequence 3767, Application US/10108260A
; Publication No. US20040005560A1
; GENERAL INFORMATION:
; APPLICANT: HELIX RESEARCH INSTITUTE
; TITLE OF INVENTION: NO. US20040005560A1el full length cDNA
; FILE REFERENCE: HI-A0106
; CURRENT APPLICATION NUMBER: US/10/108,260A
; CURRENT FILING DATE: 2002-03-27
; NUMBER OF SEQ ID NOS: 5458
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3767
; LENGTH: 519
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-108-260A-3767

Query Match
Best Local Similarity 92.9%; Score 39; DB 15; Length 519;
Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 CXCRPH 7
Db 34 CSCRPH 40

RESULT 4
US-09-753-139C-1
; Sequence 1, Application US/09753139C
; Publication No. US20030073808A1
; GENERAL INFORMATION:
; APPLICANT: Quirk, Stephen
; APPLICANT: Tyrrell, David
; TITLE OF INVENTION: Design and Use of Advanced Zinc Chelating Peptides to Regulate Me
; FILE REFERENCE: 44039-227522 11301-0200
; CURRENT APPLICATION NUMBER: US/09/753,139C
; CURRENT FILING DATE: 2000-12-29
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 7
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
1
```

```
; OTHER INFORMATION: Synthetic Peptide
; FEATURE:
; NAME/KEY: MISC FEATURE
; LOCATION: (2)..(2)
; OTHER INFORMATION: X = Ser or Thr
; FEATURE:
; NAME/KEY: MISC FEATURE
; LOCATION: (4)..(4)
; OTHER INFORMATION: X = Ser, Ala or Val
US-09-753-139C-1

Query Match
Best Local Similarity 90.5%; Score 38; DB 10; Length 7;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CXCRPH 7
Db 1 CXCRPH 7

RESULT 5
US-10-264-237-1473
; Sequence 1473, Application US/10264237
; Publication No. US20040009491A1
; GENERAL INFORMATION:
; APPLICANT: Birse et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PA331P1
; CURRENT APPLICATION NUMBER: US/10/264,237
; CURRENT FILING DATE: 2002-10-04
; PRIOR APPLICATION NUMBER: PCT/US01/16450
; PRIOR FILING DATE: 2001-05-18
; PRIOR APPLICATION NUMBER: US 60/205,515
; PRIOR FILING DATE: 2000-05-19
; NUMBER OF SEQ ID NOS: 2876
; SOFTWARE: PatentIn Ver. 3.1
; SEQ ID NO 1473
; LENGTH: 170
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: MISC FEATURE
; LOCATION: (46)
; OTHER INFORMATION: Xaa equals any of the twenty naturally occurring L-amino acids
; NAME/KEY: MISC FEATURE
; LOCATION: (126)
; OTHER INFORMATION: Xaa equals any of the twenty naturally occurring L-amino acids
US-10-264-237-1473

Query Match
Best Local Similarity 88.1%; Score 37; DB 15; Length 170;
Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 CXCRPH 7
Db 114 CLCRPH 120

RESULT 6
US-10-412-699B-1681
; Sequence 1681, Application US/10412699B
; Publication No. US20040045049A1
; GENERAL INFORMATION:
; APPLICANT: Mendel Biotechnology, Inc.
; APPLICANT: Zhang, James
; APPLICANT: Fromm, Michael E.
; APPLICANT: Heard, Jacqueline B.
; APPLICANT: Riechmann, Jose Luis
; APPLICANT: Adam, Luc J.
; APPLICANT: Broun, Pierre E.
; APPLICANT: Pineda, Omaira
; APPLICANT: Reuber, T. Lynne
```



```

FILE REFERENCE: MEI-0047 CIP
CURRENT APPLICATION NUMBER: US/10/374,780A
CURRENT FILING DATE: 2003-02-25
PRIOR APPLICATION NUMBER: 09/837,944
PRIOR FILING DATE: 2001-04-18
PRIOR APPLICATION NUMBER: 60/310,847
PRIOR FILING DATE: 2001-08-09
PRIOR APPLICATION NUMBER: 09/934,455
PRIOR FILING DATE: 2001-08-22
PRIOR APPLICATION NUMBER: 60/336,049
PRIOR FILING DATE: 2001-11-19
PRIOR APPLICATION NUMBER: 60/338,692
PRIOR FILING DATE: 2001-12-11
PRIOR APPLICATION NUMBER: 10/171,468
PRIOR FILING DATE: 2002-06-14
PRIOR APPLICATION NUMBER: 10/225,066
PRIOR FILING DATE: 2002-08-09
PRIOR APPLICATION NUMBER: 10/225,067
PRIOR FILING DATE: 2002-08-09
PRIOR APPLICATION NUMBER: 10/225,068
PRIOR FILING DATE: 2002-08-09
NUMBER OF SEQ ID NOS: 2906
SOFTWARE: PatentIn version 3.2
SEQ ID NO 1670
LENGTH: 289
TYPE: PRT
ORGANISM: Oryza sativa
FEATURE:
OTHER INFORMATION: Orthologous to G1540
US-10-374-780A-1670

Query Match
Best Local Similarity 81.0%; Score 34; DB 15; Length 289;
Matches 4; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 CXCXPH 7
| | | |
| | | |
Db 279 CSCSPY 285

RESULT 8
US-10-424-599-153612
Sequence 153612, Application US/10424599
Publication No. US20040031072A1
GENERAL INFORMATION:
APPLICANT: La Rosa Thomas J
APPLICANT: Kovalic David K
APPLICANT: Zhou Yihua
APPLICANT: Cao Yongwei
TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated with
FILE REFERENCE: 38-21(53223)B
CURRENT APPLICATION NUMBER: US/10/424,599
CURRENT FILING DATE: 2003-04-28
NUMBER OF SEQ ID NOS: 285684
SEQ ID NO 153612
LENGTH: 91
TYPE: PRT
ORGANISM: Glycine max
FEATURE:
OTHER INFORMATION: Clone ID: PAT_MRT3847_109735C.1.pep
US-10-424-599-153612

Query Match
Best Local Similarity 78.6%; Score 33; DB 12; Length 91;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CXCXPH 6
| | | |
| | | |
Db 16 CSCSPH 21

RESULT 9

```

```
US-09-925-297-595
; Sequence 595, Application US/09925297
; Patent No. US20020081659A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA105
; CURRENT APPLICATION NUMBER: US/09/925,297
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05989
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 928
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 595
; LENGTH: 97
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (38)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-925-297-595

Query Match          78.6%; Score 33; DB 9; Length 97;
Best Local Similarity 66.7%; Pred. No. 4.9e+02;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY      1 CXCXPH 6
      ||||
Db      46 CACSPH 51

RESULT 10
US-09-796-692-1071
; Sequence 1071, Application US/09796692
; Publication No. US20020198362A1
; GENERAL INFORMATION:
; APPLICANT: Gaiger, Alexander
; APPLICANT: Mannion, Jane
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DETECTION, DIAGNOSIS AND THERAPY
; FILE REFERENCE: 2077.001200
; CURRENT APPLICATION NUMBER: US/09/796,692
; CURRENT FILING DATE: 2001-03-01
; PRIOR APPLICATION NUMBER: 60/186,126
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: 60/190,479
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: 60/200,545
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: 60/200,303
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: 60/200,779
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: 60/200,999
; PRIOR FILING DATE: 2000-05-01
; PRIOR APPLICATION NUMBER: 60/202,084
; PRIOR FILING DATE: 2000-05-04
; PRIOR APPLICATION NUMBER: 60/206,201
; PRIOR FILING DATE: 2000-05-22
; PRIOR APPLICATION NUMBER: 60/218,950
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: 60/222,903
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: 60/223,416
; PRIOR FILING DATE: 2000-08-04
; PRIOR APPLICATION NUMBER: 60/223,378
; PRIOR FILING DATE: 2000-08-07
; NUMBER OF SEQ ID NOS: 9597
; SOFTWARE: FastSeq for Windows Version 3.0
```

```
; SEQ ID NO 1071
; LENGTH: 108
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-796-692-1071

Query Match          78.6%; Score 33; DB 9; Length 108;
Best Local Similarity 66.7%; Pred. No. 5.3e+02;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY      1 CXCXPH 6
      ||||
Db      6 CACSPH 11

RESULT 11
US-10-040-862-1071
; Sequence 1071, Application US/10040862
; Publication No. US2003007836A1
; GENERAL INFORMATION:
; APPLICANT: Gaiger, Alexander
; APPLICANT: Mannion, Jane
; APPLICANT: Retter, Marc
; APPLICANT: Corixa Corporation
; TITLE OF INVENTION: Compositions and Methods for the Detection, Diagnosis and Therapy
; FILE REFERENCE: 014058-013520US
; CURRENT APPLICATION NUMBER: US/10/040,862
; CURRENT FILING DATE: 2001-11-06
; PRIOR APPLICATION NUMBER: US 60/186,126
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: US 60/190,479
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: US 60/200,545
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: US 60/200,303
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,779
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,999
; PRIOR FILING DATE: 2000-05-01
; PRIOR APPLICATION NUMBER: US 60/202,084
; PRIOR FILING DATE: 2000-05-04
; PRIOR APPLICATION NUMBER: US 60/206,201
; PRIOR FILING DATE: 2000-05-22
; PRIOR APPLICATION NUMBER: US 60/218,950
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 60/222,903
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: US 60/223,416
; PRIOR FILING DATE: 2000-08-04
; PRIOR APPLICATION NUMBER: US 60/223,378
; PRIOR FILING DATE: 2000-08-07
; PRIOR APPLICATION NUMBER: US 09/796,692
; NUMBER OF SEQ ID NOS: 10467
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1071
; LENGTH: 108
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-040-862-1071

Query Match          78.6%; Score 33; DB 14; Length 108;
Best Local Similarity 66.7%; Pred. No. 5.3e+02;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
```

```
RESULT 12
US-10-057-475B-1071
; Sequence 1071, Application US/10057475B
; Publication No. US20040002068A1
; GENERAL INFORMATION:
; APPLICANT: Gaiger, Alexander
; APPLICANT: Algate, Paul A.
; APPLICANT: Mannion, Jane
; APPLICANT: Clapper, Jonathan David
; APPLICANT: Wang, Aijun
; APPLICANT: Ordonez, Nadia
; APPLICANT: Carter, Lauren
; APPLICANT: McNeill, Patricia Dianne
; APPLICANT: Corixa Corporation
; TITLE OF INVENTION: Compositions and Methods for the Detection, Diagnosis and Therapy
; FILE REFERENCE: 014058-014402US
; CURRENT FILING DATE: 2002-01-22
; PRIOR APPLICATION NUMBER: US/10/057,475B
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: US 60/190,479
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: US 60/200,545
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: US 60/200,303
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,779
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,999
; PRIOR FILING DATE: 2000-05-01
; PRIOR APPLICATION NUMBER: US 60/202,084
; PRIOR FILING DATE: 2000-05-04
; PRIOR APPLICATION NUMBER: US 60/206,201
; PRIOR FILING DATE: 2000-05-22
; PRIOR APPLICATION NUMBER: US 60/218,950
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 60/222,903
; PRIOR FILING DATE: 2000-08-03
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 10979
; SOFTWARE: FASTSEQ for Windows Version 3.0
; SEQ ID NO 1071
; LENGTH: 108
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-057-475B-1071

Query Match          78.6%; Score 33; DB 15; Length 108;
Best Local Similarity 66.7%; Pred. No. 5.3e+02;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy      1 CXCPH 6
        |||
Db      6 CSCSPH 11

RESULT 13
US-10-154-884B-1071
; Sequence 1071, Application US/10154884B
; Publication No. US20040005561A1
; GENERAL INFORMATION:
; APPLICANT: Gaiger, Alexander
; APPLICANT: Algate, Paul A.
; APPLICANT: Mannion, Jane
; APPLICANT: Retter, Marc W.
; APPLICANT: Corixa Corporation
; TITLE OF INVENTION: Compositions and Methods for the Detection, Diagnosis and Therapy
; FILE REFERENCE: 014058-01521US
; CURRENT FILING DATE: US/10/154,884B
; CURRENT FILING DATE: 2002-05-23
; PRIOR APPLICATION NUMBER: US 60/186,126
```

```
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: US 60/190,479
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: US 60/200,545
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: US 60/200,303
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,779
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,999
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,999
; PRIOR FILING DATE: 2000-05-01
; PRIOR APPLICATION NUMBER: US 60/202,084
; PRIOR FILING DATE: 2000-05-04
; PRIOR APPLICATION NUMBER: US 60/206,201
; PRIOR FILING DATE: 2000-05-22
; PRIOR APPLICATION NUMBER: US 60/218,950
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 60/222,903
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: US 60/223,416
; PRIOR FILING DATE: 2000-08-04
; PRIOR APPLICATION NUMBER: US 60/223,378

US-10-154-884B-1071

Query Match          78.6%; Score 33; DB 15; Length 108;
Best Local Similarity 66.7%; Pred. No. 5.3e+02;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy      1 CXCPH 6
        |||
Db      6 CSCSPH 11

RESULT 14
US-09-796-692-1736
; Sequence 1736, Application US/09796692
; Publication No. US20020198362A1
; GENERAL INFORMATION:
; APPLICANT: Gaiger, Alexander
; APPLICANT: Mannion, Jane
; APPLICANT: Corixa Corporation
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DETECTION, DIAGNOSIS AND THERAPY
; FILE REFERENCE: 2077.001200
; CURRENT APPLICATION NUMBER: US/09/796,692
; CURRENT FILING DATE: 2001-03-01
; PRIOR APPLICATION NUMBER: 60/186,126
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: 60/190,479
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: 60/200,545
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: 60/200,303
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: 60/200,779
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: 60/200,999
; PRIOR FILING DATE: 2000-05-01
; PRIOR APPLICATION NUMBER: 60/202,084
; PRIOR FILING DATE: 2000-05-04
; PRIOR APPLICATION NUMBER: 60/206,201
; PRIOR FILING DATE: 2000-05-22
; PRIOR APPLICATION NUMBER: 60/218,950
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: 60/222,903
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: 60/223,416
; PRIOR FILING DATE: 2000-08-04
; PRIOR APPLICATION NUMBER: 60/223,378
```

; PRIOR FILING DATE: 2000-08-07
 ; NUMBER OF SEQ ID NOS: 9597
 ; SOFTWARE: FastSEQ for Windows Version 3.0
 ; SEQ ID NO 1736
 ; LENGTH: 119
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: variant
 ; LOCATION: (1)...(119)
 ; OTHER INFORMATION: Xaa = Any amino acid
 US-09-796-692-1736

Query Match 78.6%; Score 33; DB 9; Length 119;
 Best Local Similarity 66.7%; Pred. No. 5.7e+02;
 Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CXCKPH 6
 Db 6 CSCSPH 11

RESULT 15
 US-10-040-862-1736
 ; Sequence 1736, Application US/10040862
 ; Publication No. US20030078396a1
 ; GENERAL INFORMATION:
 ; APPLICANT: Gaiger, Alexander
 ; APPLICANT: Algate, Paul A.
 ; APPLICANT: Mannion, Jane
 ; APPLICANT: Retter, Marc
 ; APPLICANT: Corixa Corporation
 ; TITLE OF INVENTION: Compositions and Methods for the Detection, Diagnosis and Therapy
 ; FILE REFERENCE: 014058-013520US
 ; CURRENT APPLICATION NUMBER: US/10/040, 862
 ; CURRENT FILING DATE: 2001-11-06
 ; PRIOR APPLICATION NUMBER: US 60/186,126
 ; PRIOR FILING DATE: 2000-03-01
 ; PRIOR APPLICATION NUMBER: US 60/190,479
 ; PRIOR FILING DATE: 2000-03-17
 ; PRIOR APPLICATION NUMBER: US 60/200,545
 ; PRIOR FILING DATE: 2000-04-27
 ; PRIOR APPLICATION NUMBER: US 60/200,303
 ; PRIOR FILING DATE: 2000-04-28
 ; PRIOR APPLICATION NUMBER: US 60/200,779
 ; PRIOR FILING DATE: 2000-04-28
 ; PRIOR APPLICATION NUMBER: US 60/200,999
 ; PRIOR FILING DATE: 2000-05-01
 ; PRIOR APPLICATION NUMBER: US 60/202,084
 ; PRIOR FILING DATE: 2000-05-04
 ; PRIOR APPLICATION NUMBER: US 60/206,201
 ; PRIOR FILING DATE: 2000-05-22
 ; PRIOR APPLICATION NUMBER: US 60/218,950
 ; PRIOR FILING DATE: 2000-07-14
 ; PRIOR APPLICATION NUMBER: US 60/222,903
 ; PRIOR FILING DATE: 2000-08-03
 ; PRIOR APPLICATION NUMBER: US 60/223,416
 ; PRIOR FILING DATE: 2000-08-04
 ; PRIOR APPLICATION NUMBER: US 60/223,378
 ; PRIOR FILING DATE: 2000-08-07
 ; PRIOR APPLICATION NUMBER: US 09/796,692
 ; PRIOR FILING DATE: 2001-03-01
 ; NUMBER OF SEQ ID NOS: 10467
 ; SOFTWARE: FastSEQ for Windows Version 3.0
 ; SEQ ID NO 1736
 ; LENGTH: 119
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: variant
 ; LOCATION: (1)...(119)
 ; OTHER INFORMATION: Xaa = Any amino acid

US-10-040-862-1736

Query Match 78.6%; Score 33; DB 14; Length 119;
 Best Local Similarity 66.7%; Pred. No. 5.7e+02;
 Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CXCKPH 6
 Db 6 CSCSPH 11

Search completed: April 8, 2004, 11:55:41
 Job time : 30.3333 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using SW model

Run on: April 8, 2004, 11:00:44 ; Search time 44.0417 Seconds
(without alignments)
44.908 Million cell updates/sec

Title: US-09-753-139c-1

Perfect score: 42
Sequence: 1 CKCXPMP 7

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database :

A_Geneseq_29Jan04:*
1: geneseq1980s:*
2: geneseq1990s:*
3: geneseq2000s:*
4: geneseq2001s:*
5: geneseq2002s:*
6: geneseq2003as:*
7: geneseq2003bs:*
8: geneseq2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	40	95.2	55	4 AAU14695	AAU14695 Novel bon
2	40	95.2	70	5 ABP06775	ABP06775 Human ORF
3	39	92.5	30	5 ABP79208	ABP79208 Human pro
4	38	90.5	7	5 ABH83462	ABH83462 Tissue in
5	38	90.5	73	4 AAU40466	AAU40466 Propionib
6	38	90.5	73	6 ABM36985	ABM36985 Propionib
7	38	90.5	93	5 ABP58502	ABP58502 Human sho
8	38	90.5	119	4 AAU06377	AAU06377 Human foe
9	38	90.5	221	4 AAH93924	AAH93924 Human pro
10	38	90.5	225	3 AAU95940	AAU95940 Porcine a
11	38	90.5	527	6 ABP72624	ABP72624 Porcine chl
12	37	88.1	122	4 AAU29832	AAU29832 Novel hum
13	37	88.1	170	5 ABH89097	ABH89097 Human pol
14	34	81.0	104	4 AAU49509	AAU49509 Propionib
15	34	81.0	104	6 ABM46028	ABM46028 Propionib
16	33	78.6	36	3 AAB44849	AAB44849 Human sec
17	33	78.6	86	4 ABG23045	ABG23045 Novel hum
18	33	78.6	97	3 AAB54143	AAB54143 Human pan
19	33	78.6	108	4 AAU80707	AAU80707 Human hae
20	33	78.6	119	4 AAU81372	AAU81372 Human hae
21	33	78.6	120	4 AAU80701	AAU80701 Human hae
22	33	78.6	139	4 AAU80391	AAU80391 Human hae
23	33	78.6	149	4 AAU80713	AAU80713 Human hae
24	33	78.6	155	4 AAU81278	AAU81278 Human hae
25	33	78.6	592	7 ADE86252	ADE86252 Hedgehog

ALIGNMENTS

26	32	76.2	40	2	AAU00299	AAU00299 Human sec
27	32	76.2	40	7	ADD90328	ADD90328 Novel hum
28	32	76.2	64	4	AAU51994	AAU51994 Propionib
29	32	76.2	64	6	ABM48513	ABM48513 Propionib
30	32	76.2	80	5	ABP02671	ABP02671 Human ORF
31	32	76.2	94	4	ABG11465	ABG11465 Novel hum
32	32	76.2	110	4	ABH17019	ABH17019 Human hae
33	32	76.2	117	6	ABP76241	ABP76241 Human GEN
34	32	76.2	118	4	ABG24883	ABG24883 Novel hum
35	32	76.2	125	4	ABG14479	ABG14479 Novel hum
36	32	76.2	129	3	AAU03981	AAU03981 Human sec
37	32	76.2	134	5	ABU00091	ABU00091 Human nov
38	32	76.2	135	5	ABR01728	ABR01728 Human bre
39	32	76.2	148	4	ABG19243	ABG19243 Novel hum
40	32	76.2	149	3	AAU54086	AAU54086 Human pan
41	32	76.2	164	4	AAU75009	AAU75009 Human col
42	32	76.2	171	4	ABG20830	ABG20830 Novel hum
43	32	76.2	194	7	ADB64752	ADB64752 Human pro
44	32	76.2	215	4	AAU63957	AAU63957 Human pro
45	32	76.2	215	4	ABG13901	ABG13901 Novel hum

RESULT 1
AAU14695
AAU14695 standard; protein; 55 AA.

AAU14695;
24-OCT-2001 (first entry)

Novel bone marrow polypeptide #94.

Bone marrow; diagnostic; therapeutic; gene therapy; antigenic;
haematopoiesis; myeloid; lymph cell disorder; tissue regeneration;
wound healing; nutritional supplement; immune disorder;
severe combined immunodeficiency; SCID.

Homo sapiens.

MO200157187-A2.

09-AUG-2001.

05-FEB-2001; 2001WO-US003782.

03-FEB-2000; 2000US-0046914.
PR 20-JUN-2000; 2000US-00588075.
PR 19-JUL-2000; 2000US-00620325.
PR 30-NOV-2000; 2000US-0250683P.

(HYSE-) HYSEQ INC.

Ford JR, Boyle BJ, Tang YT, Liu C, Asundi V, Zhou P, Xue AJ;
Ren F, Drmanac RT;
WPI; 2001-488875/53.
N-PSDB; AAS23000.

Nucleic acids encoding bone marrow polypeptides, useful in diagnostic and gene therapy.

Claim 10; Page 251-252; 392pp; English.

AAU14602-AAU14794 represent novel bone marrow polypeptides of the invention. The proteins and corresponding coding sequences may be used in the prevention, diagnosis and treatment of diseases associated with inappropriate bone marrow polypeptide expression. For example, to treat disorders associated with decreased expression by rectifying mutations or deletions in a patient's genome that affect the activity of the polypeptides by expressing inactive proteins or to supplement the

CC patient's own production of the polypeptide. Additionally, the nucleic
 CC acids may be used to produce the polypeptides, by inserting the nucleic
 CC acids into a host cell and culturing the cell to express the protein. The
 CC nucleic acid and its complementary sequences may also be used as DNA
 CC probes in diagnostic assays to detect and quantitate the presence of
 CC similar nucleic acid sequences in samples, and therefore which patients
 CC may be in need of restorative therapy. The proteins may also be used as
 CC antigens in the production of antibodies against bone marrow proteins and
 CC in assays to identify modulators of their expression and activity. The
 CC anti-bone marrow protein antibodies and antagonists may also be used to
 CC down regulate expression and activity. The antibodies may also be used as
 CC diagnostic agents for detecting the presence of the protein in samples
 CC (e.g. by enzyme linked immunosorbent assay (ELISA)). The proteins may be
 CC used to regulate haematopoiesis activity, and consequently in the
 CC treatment of myeloid or lymph cell disorders; in tissue regeneration,
 CC such as wound healing; as a nutritional supplement; and in treatment of
 CC immune disorders such as severe combined immunodeficiency (SCID)

XX SQ Sequence 55 AA;

Query Match 95.2%; Score 40; DB 4; Length 55;
 Best Local Similarity 71.4%; Pred. No. 41;
 Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 CXCXPHP 7
 | | | | |
 Db 10 CXCXPHP 16

RESULT 2
 ABP06775
 ID ABP06775 standard; protein; 70 AA.
 AC ABP06775;
 XX
 XX 24-JUN-2002 (first entry)
 DT
 XX
 XX
 DE Human ORFX protein sequence SEQ ID NO:13532.

XX Human, open reading frame; ORFX; gene therapy; cancer; cirrhosis;
 KM hyperproliferative disorder; psoriasis; benign tumour; hemorrhage;
 KM degenerative disorder; osteoarthritis; neurodegenerative disorder;
 KM cardiovascular disease; diabetes mellitus; systemic lupus erythematosus;
 KM hypertension; hypothyroidism; cholesterol ester storage disease;
 KM immune deficiency; immune disorder; infectious disease;
 KM autoimmune disorder; rheumatoid arthritis; autoimmune thyroiditis;
 KM myasthenia gravis.

XX Homo sapiens.

XX WO200192523-A2.

XX 06-DEC-2001.

XX 29-MAY-2001; 2001WO-US010836.

XX 30-MAY-2000; 2000US-0206132P.

XX 29-AUG-2000; 2000US-0228716P.

XX (CURA-) CURAGEN CORP.

XX Shinketsu RA, Leach MD;

XX WPI, 2002-106308/14.

XX N-PSDB; ABN22527.

XX Novel human polypeptides and polynucleotides useful for diagnosing,
 PT preventing and treating cardiovascular disease, neurodegenerative,
 PT hyperproliferative disorders and autoimmune disorders.

XX Disclosure, SEQ ID NO 13532; 1037pp; English.

XX The present invention describes substantially purified human proteins

CC (referred to as open reading frame, ORFX, where X is 1-11491 (see Table 1
 CC in the specification). ABN15762 to ABN27252 encode the human ORFX
 CC proteins given in ABP00010 to ABP11500. ORFX proteins are useful for
 CC treating or preventing a pathology associated with an ORFX-associated
 CC disorder in humans, and in the manufacture of a medicament for treating a
 CC syndrome associated with ORFX-associated disorder. ORFX polynucleotide
 CC sequences can be used in gene therapy. ORFX sequences can be used in the
 CC treatment of cancer. Hyperproliferative disorders, cirrhosis of liver,
 CC psoriasis, benign tumours, keloid, degenerative disorders, hemorrhage,
 CC osteoarthritis, neurodegenerative disorders, disorders related to organ
 CC transplantation, cardiovascular diseases, diabetes mellitus, systemic
 CC lupus erythematosus, hypertension, hypothyroidism, cholesterol ester
 CC storage disease, various immune deficiencies and disorders, infectious
 CC diseases, autoimmune disorders such as multiple sclerosis, rheumatoid
 CC arthritis, autoimmune thyroiditis, myasthenia gravis, graft-versus-host
 CC disease and autoimmune inflammatory eye disease. ORFX proteins are also
 CC useful for treating burns, incisions, ulcers, for treating osteoporosis,
 CC bone degenerative disorders, or periodontal disease, and for gut
 CC protection or regeneration and treatment of lung or liver fibrosis,
 CC reperfusion injury in various tissues and conditions resulting from
 CC systemic cytokine damage. N.B. The sequence data for this patent did not
 CC form part of the printed specification, but was obtained in electronic
 CC format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
 XX

SQ Sequence 70 AA;

Query Match 95.2%; Score 40; DB 5; Length 70;
 Best Local Similarity 71.4%; Pred. No. 50;
 Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 CXCXPHP 7
 | | | | |
 Db 49 CXCXPHP 55

RESULT 3
 ABB79208
 ID ABB79208 standard; protein; 30 AA.
 XX
 XX ABB79208;
 AC
 XX 08-AUG-2002 (first entry)
 DT
 XX
 XX
 DE Human prostate specific protein sequence SEQ ID NO:157.

XX Human prostate specific protein sequence SEQ ID NO:157.
 KM Human, prostate specific gene; prostate specific protein; PSG; PSP;
 KM prostate cancer.

XX Homo sapiens.

XX WO200236808-A2.

XX 10-MAY-2002.

XX 05-NOV-2001; 2001WO-US047283.

XX 03-NOV-2000; 2000US-0245740P.

XX (DIAD-) DIADEXUS INC.

XX Sun Y, Reipon H, Chen S, Liu C;

XX WPI, 2002-471506/50.

XX New prostate-specific nucleic acids and polypeptides, useful for
 PT identifying, diagnosing, monitoring, staging, imaging, and treating
 PT prostate cancer and non-cancerous disease states in prostate tissue.

XX Claim 11; Page 223; 254pp; English.

XX ABN87650 to ABN87789 represent human prostate-specific nucleic acids (I),
 CC and ABB79192 to ABB79295 represent human prostate-specific proteins (II)
 CC from the present invention. (I) and (II) have cytostatic activity. (I)

CC can be used in gene therapy. The prostate-specific nucleic acids.
CC polypeptides and compositions from the present invention can be used for
CC identifying, diagnosing, monitoring, staging, imaging, and treating
CC prostate cancer and non-cancerous disease states in prostate tissue; for
CC identifying prostate tissue; for monitoring, identifying and/or designing
CC agonists and antagonists of the polypeptides; in gene therapy; in
CC producing transgenic animals and cells; for producing engineered prostate
CC tissue for treatment and research; and as elements in an array or
CC computer program for pattern recognition of prostate disorders. The
CC nucleic acids may be used as hybridisation probes to detect, characterise
CC and quantify hybridising nucleic acids in, and isolate hybridising
CC nucleic acids from, both genomic and transcript-derived nucleic acid
CC samples
CC
SQ Sequence 30 AA;
Query Match 92.9%; Score 39; DB 5; Length 30;
Best Local Similarity 71.4%; Pred. No. 36;
Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 CXCXPHP 7
DB 24 CACFPHP 30
RESULT 4
ABB83462
ID ABB83462 standard; peptide; 7 AA.
XX ABB83462;
XX
XX 30-SEP-2002 (first entry)
XX
XX Tissue Inhibitor of Metalloproteinase, TIMP, derived peptide #1.
XX
XX MMP; Matrix Metalloproteinase; zinc chelator; chronic wound; acute wound;
XX Tissue Inhibitor of Metalloproteinase; TIMP; connective tissue breakdown;
XX angiogenesis-associated disorder.
XX
XX Synthetic.
XX
XX Key Location/Qualifiers
XX Misc-difference 2
XX PT /label= Ser, Thr
XX FT Misc-difference 4
XX FT /label= Ser, Ala, Val
XX
XX WO200253173-A2.
XX
XX PD 11-JUL-2002.
XX
XX PF 21-DEC-2001; 2001WO-US049276.
XX
XX PR 29-DEC-2000; 2000US-00753139.
XX
XX (KIMB) KIMBERLY-CLARK WORLDWIDE INC.
XX
XX PI Quirk S, Tyrrell DJ;
XX
XX WPI; 2002-583595/62.
XX
XX PT New matrix metalloproteinase regulator useful for the treatment of
XX PT chronic and acute wounds comprises a zinc chelator and a tissue inhibitor
XX PT of metalloproteinase-derived peptide.
XX
XX PS Claim 5; Page 15; 57pp; English.
XX
XX CC The present invention relates to Matrix Metalloproteinase (MMP)
XX CC regulators, which comprise a zinc chelator and a tissue inhibitor of
XX CC Metalloproteinases (TIMP)-derived peptide. The present sequence is one
XX CC such TIMP-derived peptide used to generate the MMP regulators. The MMP
XX CC regulators are useful for treating chronic and acute wounds, angiogenesis
XX CC -associated disorders, and other diseases and disorders involving

CC uncontrolled breakdown of connective tissues by MMPs. MMPs contain a zinc
CC molecule located in the active site, which participates in degrading
CC collagen. The binding specificity of the TIMP-derived peptide brings the
CC zinc chelator into molecular proximity of the MMP bound zinc in such a
CC way to allow ligation. This results in the regulation of the level of MMP
CC activity to promote wound healing by providing a MMP regulator having
CC high affinity and selectivity
CC
SQ Sequence 7 AA;
Query Match 90.5%; Score 38; DB 5; Length 7;
Best Local Similarity 100.0%; Pred. No. 1.4e+06;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CXCXPHP 7
DB 1 CXCXPHP 7
RESULT 5
AAU40466
ID AAU40466 standard; protein; 73 AA.
XX AAU40466;
XX
XX 13-FEB-2002 (first entry)
XX
XX Propionibacterium acnes immunogenic protein #1362.
XX
XX SAPHO syndrome; synovitis; acne; pustulosis; hyperostosis; osteomyelitis;
XX uveitis; endophthalmitis; bone; joint; central nervous system; ELISA;
XX inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay;
XX dermatological; osteopathic; neuroprotectant.
XX
XX OS Propionibacterium acnes.
XX
XX PN WO200101581-A2.
XX
XX PD 01-NOV-2001.
XX
XX PF 20-APR-2001; 2001WO-US012865.
XX
XX PR 21-APR-2000; 2000US-0199047P.
XX PR 02-JUN-2000; 2000US-0208841P.
XX PR 07-JUL-2000; 2000US-0216747P.
XX
XX PA (CORI-) CORIXA CORP.
XX
XX PI Skelky YAW, Persing DH, Mitcham JL, Wang SS, Bhatia A;
XX PI L'maisonneuve J, Zhang Y, Yen S, Carter D;
XX
XX DR WPI; 2001-616774/71.
XX
XX DR N-PSDB; AAS59511.
XX
XX PT Propionibacterium acnes polypeptides and nucleic acids useful for
XX PT vaccinating against and diagnosing infections, especially useful for
XX PT treating acne vulgaris.
XX
XX PS Example 1; SEQ ID NO 1661; 1069pp; English.
XX
XX CC Sequences AAU39105-AAU6017 represent Propionibacterium acnes immunogenic
XX CC polypeptides. The proteins and their associated DNA sequences are used in
XX CC the treatment, prevention and diagnosis of medical conditions caused by
XX CC P. acnes. The disorders include SAPHO syndrome (synovitis, acne,
XX CC pustulosis, hyperostosis and osteomyelitis), uveitis and endophthalmitis.
XX CC P. acnes is also involved in infections of bone, joints and the central
XX CC nervous system, however it is particularly involved in the inflammatory
XX CC lesions associated with acne vulgaris. A method for detecting the
XX CC presence or absence of P. acnes in a patient comprises contacting a
XX CC sample with a binding agent that binds to the proteins of the invention
XX CC and determining the amount of bound protein in the sample. The
XX CC polypeptides may be used as antigens in the production of antibodies
XX CC specific for P. acnes proteins. These antibodies can be used to

CC downregulate expression and activity of *P. acnes* polypeptides and
 CC therefore treat *P. acnes* infections. The antibodies may also be used as
 CC diagnostic agents for determining *P. acnes* presence, for example, by
 CC enzyme linked immunosorbent assay (ELISA). Note: The sequence data for
 CC this patent did not form part of the printed specification, but was
 CC obtained in electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences

XX
 SQ Sequence 73 AA;

Query Match 90.5%; Score 38; DB 4; Length 73;
 Best Local Similarity 71.4%; Pred. No. 1e+02;

Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CXCKP 7
 DB 56 CRCIRPH 62

RESULT 6
 ABM36985
 ID ABM36985 standard; protein, 73 AA.
 XX
 AC ABM36985;
 XX
 DT 20-OCT-2003 (first entry)
 XX
 DE Propionibacterium acnes predicted ORF-encoded polypeptide #1661.

XX Acne vulgaris; antiacne; dermatological; antibacterial;
 KW immunostimulant; immune response; vaccine.
 XX
 OS Propionibacterium acnes.
 XX
 PN WO200303515-A1.
 XX
 PD 24-APR-2003.
 XX
 PE 11-OCT-2002; 2002WO-US032727.
 XX
 PR 15-OCT-2001; 2001US-00978925.
 XX
 PA (CORI-) CORIXA CORP.

XX Mitcham JI, Skeiky YAW, Persing DH, Bhatia A, Maisonneuve JI;
 PI Zhang Y, Wang S, Jen S, Lodes MJ, Benson DR, Jones R, Carter D;
 PI Barth B, Vallie-Douglas J;
 XX
 DR WPI; 2003-381789/36.
 DR N-PSDB; ACF64440.

XX New Propionibacterium acnes polypeptides and polynucleotides encoding the
 PT polypeptide, useful for diagnosing, preventing or treating acne vulgaris,
 PT or for stimulating an immune response specific for *P. acnes* protein.
 PS
 XX Example 1; SEQ ID NO 1661; 1481bp; English.

XX The invention relates to an isolated polynucleotide (ACF64435-ACF64733)
 CC encoding a Propionibacterium acnes protein. The invention also relates to
 CC polypeptides encoded by the polynucleotides (ABM36985-ABM64536) and to
 CC immunogenic fragments of *P. acnes* polypeptides. The invention
 CC additionally encompasses expression vectors and host cells comprising a
 CC polynucleotide of the invention; antibodies against polypeptides of the
 CC invention; fusion proteins comprising a polypeptide of the invention; a
 CC method for stimulating an immune response specific for *P. acnes*
 CC polypeptide and an isolated T cell population comprising T cells prepared
 CC via this method; a vaccine composition (comprising *P. acnes* polypeptides,
 CC polynucleotides, antibodies, fusion proteins, T cell populations, or
 CC antigen-presenting cells that express the polypeptide); a method and kit
 CC for detecting or determining the presence or absence of *P. acnes* in a
 CC patient; and a method for inhibiting the development of *P. acnes* in a
 CC patient. The *P. acnes* polypeptides, polynucleotides, antibodies, fusion
 CC proteins, T cell populations or antigen-presenting cells that express the

CC polypeptides are useful for diagnosing, preventing or treating acne
 CC vulgaris, or for stimulating an immune response specific for *P. acnes*
 CC protein. The polynucleotides can also be used as probes or primers for
 CC nucleic acid hybridisation. The vaccine composition is useful for the
 CC stimulation of an immune response against *P. acnes*, or for treating acne,
 CC and the kit is useful for performing a diagnostic assay. The present
 CC sequence represents a polypeptide predicted to be encoded by an ORF (open
 CC reading frame) contained within the *P. acnes* polynucleotides of the
 CC invention. Note: The sequence data for this patent did not form part of
 CC the printed specification, but was obtained in electronic format directly
 CC from WIPO at ftp.wipo.int/pub/published_pct_sequences

XX
 SQ Sequence 73 AA;

Query Match 90.5%; Score 38; DB 6; Length 73;
 Best Local Similarity 71.4%; Pred. No. 1e+02;

Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CXCKP 7
 DB 56 CRCIRPH 62

RESULT 7
 ABP58502
 ID ABP58502 standard; protein, 93 AA.
 XX
 AC ABP58502;
 XX
 DT 20-FEB-2003 (first entry)
 XX
 DE Human short chain dehydrogenase 10.23.

XX Human; short chain dehydrogenase 10.23; recombinant production;
 KW gene therapy; malignant tumour; cancer; blood diseases;
 KW human immunodeficiency virus; HIV infection; immune disorder;
 KW inflammatory condition; cytostatic; antiinflammatory; immunomodulator;
 KW enzyme.
 XX
 OS Homo sapiens.
 XX
 PN CN1358846-A.
 XX
 PD 17-JUL-2002.
 XX
 PE 13-DEC-2000; 2000CN-00127866.
 XX
 PR 13-DEC-2000; 2000CN-00127866.
 XX
 PA (SHAN-) SHANGHAI BIOWINDOW GENE DEV INC.

XX Mao Y, Xie Y;
 PI WPI; 2002-733567/80.
 PI N-PSDB; ABV76533.
 XX
 DR Novel polypeptide-human short chain dehydrogenase 10.23 the
 PT polynucleotide for encoding the polypeptide.
 PT
 PS Claim 1; Page 26 (Disclosure); 32pp; Chinese.

XX The invention relates to human short chain dehydrogenase 10.23 (ABP58502)
 CC and nucleic acids encoding it (ABV76533). The protein has a molecular
 CC weight of 10.23 kD. The invention also relates to a method for the
 CC recombinant production of the protein, an antagonist of the protein, and
 CC the use of the protein, gene and antagonist in therapeutic applications.
 CC Short chain dehydrogenase 10.23 can be used in the treatment of a variety
 CC of diseases such as malignant tumours, blood diseases, HIV (human
 CC immunodeficiency virus) infection, immune disorders and inflammatory
 CC conditions. The present sequence represents human short chain
 CC dehydrogenase 10.23
 XX
 SQ Sequence 93 AA;

Query Match 90.5%; Score 38; DB 5; Length 93;
 Best Local Similarity 71.4%; Pred. No. 1.2e+02;
 Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CXCXPHP 7
 DB 22 CXCXPHP 28

RESULT 8
 AAM06377
 ID AAM06377 standard; protein; 119 AA.

XX AAM06377;
 AC
 XX
 DT 05-OCT-2001 (first entry)
 XX
 DE Human foetal protein, SEQ ID NO: 108.
 XX
 KW Human; foetal protein; cytotectic; immunosuppressive; immunostimulant;
 KW nootropic; neuroprotective; thrombolytic; osteopathic; antiinflammatory;
 KW gene therapy; antisense therapy; cancer; immune disorder;
 KW growth disorder; osteoporosis; thrombolytic disorder;
 KW nervous system disorder; inflammation.

XX Homo sapiens.
 XX WO200155339-A2.
 XX
 PD 02-AUG-2001.
 XX
 PF 25-JAN-2001; 2001WO-US002723.
 XX
 PR 25-JAN-2000; 2000US-00491404.
 PR 15-SEP-2000; 2000US-00663870.
 PR 06-NOV-2000; 2000US-00707351.

XX (HYSE-) HYSEQ INC.
 XX
 PI Yeung G, Ford JE, Boyle BJ, Arterburn MC, Dermanac RA, Tang YT;
 PI Liu C, Asundi V, Zhou P, Werhman T;
 XX
 DR MPI: 2001-465571/50.
 DR N-PSDB; AAH94052.
 XX
 PT Novel fetal proteins useful for the treatment and diagnosis of diseases
 PT associated with dysfunction of the protein e.g. cancers, immune
 PT disorders, growth disorders, thrombolytic disorders, nervous system
 PT disorders and inflammation.

PS Claim 10; Page 220; 715pp; English.
 PS
 CC The invention relates to novel foetal polypeptides encoded by
 CC polynucleotides comprising one of 477 sequences fully defined in the
 CC specification. The foetal polynucleotides and polypeptides are useful in
 CC the treatment and diagnosis of diseases such as cancers, immune
 CC disorders, growth disorders (e.g. osteoporosis), thrombolytic disorders,
 CC nervous system disorders and inflammation. The present sequence is a
 CC polypeptide encoded by a cDNA assembled using an expressed sequence tag
 CC (EST) found to be expressed in human foetal tissue cDNA libraries
 XX
 SQ Sequence 119 AA;

Query Match 90.5%; Score 38; DB 4; Length 119;
 Best Local Similarity 71.4%; Pred. No. 1.5e+02;
 Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CXCXPHP 7
 DB 13 CXCXPHP 19

RESULT 9
 AAB93924
 ID AAB93924 standard; protein; 221 AA.

XX AAB93924;
 AC
 XX
 DT 26-JUN-2001 (first entry)
 XX
 DE Human protein sequence SEQ ID NO:13913.

XX Human; primer; detection; diagnosis; antisense therapy; gene therapy.
 XX Homo sapiens.
 XX
 PN EP1074617-A2.
 XX
 PD 07-FEB-2001.
 XX
 PF 28-JUL-2000; 2000EP-00116126.
 XX
 PR 29-JUL-1999; 99JP-00248036.
 PR 27-AUG-1999; 99JP-00300253.
 PR 11-JAN-2000; 2000JP-00118776.
 PR 02-MAY-2000; 2000JP-00183767.
 PR 09-JUN-2000; 2000JP-00241899.

XX (HELI-) HELIX RES INST.
 XX
 PI Oca T, Isegai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
 PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
 XX
 DR MPI: 2001-318749/34.

XX
 PT Primer sets for synthesizing polynucleotides, particularly the 5602 full-
 PT length cDNAs defined in the specification, and for the detection and/or
 PT diagnosis of the abnormality of the proteins encoded by the full-length
 PT cDNAs.

PS Claim 8; SEQ ID NO 13913; 2537pp + Sequence listing; English.
 XX
 CC The present invention describes primer sets for synthesizing 5602 full-
 CC length cDNAs defined in the specification. Where a primer set comprises:
 CC (a) an oligo-dr primer and an oligonucleotide complementary to the
 CC complementary strand of a polynucleotide which comprises one of the 5602
 CC nucleotide sequences defined in the specification, where the
 CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination
 CC of an oligonucleotide comprising a sequence complementary to the
 CC complementary strand of a polynucleotide which comprises a 5'-end
 CC sequence and an oligonucleotide comprising a sequence complementary to a
 CC polynucleotide which comprises a 3'-end sequence, where the
 CC oligonucleotide comprises at least 15 nucleotides and the combination of
 CC the 5'-end sequence/3'-end sequence is selected from those defined in the
 CC specification. The primer sets can be used in antisense therapy and in
 CC gene therapy. The primers are useful for synthesizing polynucleotides,
 CC particularly full-length cDNAs. The primers are also useful for the
 CC detection and/or diagnosis of the abnormality of the proteins encoded by
 CC the full-length cDNAs. The primers allow obtaining of the full-length
 CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
 CC AAH13631 to AAH18742 represent human cDNA sequences; AAB92446 to AAB95893
 CC represent human amino acid sequences; and AAH11629 to AAH13632 represent
 CC oligonucleotides, all of which are used in the exemplification of the
 CC present invention
 XX
 SQ Sequence 221 AA;

Query Match 90.5%; Score 38; DB 4; Length 221;
 Best Local Similarity 71.4%; Pred. No. 2.4e+02;
 Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CXCXPHP 7
 DB 13 CXCXPHP 19

```

RESULT 10
AAV95940 ID AAV95940 standard; protein; 225 AA.
XX
AC AAV95940;
XX
DT 12-SEP-2003 (revised)
DT 20-NOV-2000 (first entry)
XX
DE Porcine adenovirus 3 Iva2 protein.
XX
KM PAD3; adenovirus; mastadenovirus; vector; Iva2 protein.
XX
OS Porcine adenovirus 3.
XX
PN WO200050076-A1.
XX
PD 31-AUG-2000.
XX
PF 24-FEB-2000; 2000MO-US004711.
XX
PR 24-FEB-1999; 99US-0121647P.
XX
PA (PURD ) PURDUE RES FOUND.
XX
PI Mittal SK, Aggarwal N;
XX
DR WPI; 2000-572039/53.
DR N-PESDB; AAA50439.
XX
PT Isolated nucleic acid molecules encoding proteins from porcine
PT adenovirus, used for controlling viral replication, especially in the
PT latent viral stage.
XX
PS Example 1; Fig 1G; 77pp; English.
XX
CC The present sequence is that of the protein encoded by an open reading
CC frame, Iva2, identified in a porcine adenovirus 3 (Pad3) genome fragment
CC (see AAA50439). Iva2 shows 77% and 72% amino acid identity with Iva2
CC proteins of human adenovirus 5 and bovine adenovirus 3, respectively. The
CC invention provides recombinant adenovirus vectors including novel open
CC reading frame 97R, 162R, 163R*, 288R or 184R and at least 1 foreign
CC promoter. In a preferred form of the invention, a plasmid is provided
CC that includes the nucleotide sequence encoding the novel Pad3 protein,
CC along with an adenoviral origin of replication. A replication-defective
CC Pad3 vector comprises a Pad3 genome having a functional deletion in the
CC E1 region, and with a nucleotide sequence encoding a protein of interest
CC inserted into the E1 region. Methods for making an adenovirus vector and
CC for expressing a desired protein or a viral protein in a host cell are
CC claimed. (Updated on 12-SEP-2003 to standardise OS field)
XX
SQ Sequence 225 AA;
XX
Query Match 90.5%; Score 38; DB 3; Length 225;
Best Local Similarity 71.4%; Pred. No. 2.4e+02;
Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 CXCPHP 7
DB 169 CTCPPHP 175

```

```

XX
KM Insecticide; pesticide; insect control; insect; toxin; chitinase; enzyme;
KM neuropeptide; transgenic plant; crop protection; prawn.
XX
OS Marsupeneus japonicus.
XX
PN WO2003014150-A2.
XX
PD 20-FEB-2003.
XX
PF 06-AUG-2002; 2002MO-GB003598.
XX
FR 08-AUG-2001; 2001GB-00019274.
XX
PA (UYDU-) UNIV DURHAM.
PA (ENVI-) DEPT ENVIRONMENT FOOD & RURAL AFFAIRS.
XX
PI Gatehouse JA, Fitches EC, Edwards JP;
XX
DR WPI; 2003-278469/27.
XX
PT Fusion protein useful for combating insect pests, comprises a
PT translocating moiety comprising a plant protein capable of acting as a
PT carrier to translocate toxic moiety inside plant pathogen, and a toxic
PT moiety.
XX
PS Claim 7; Fig 7; 51pp; English.
XX
CC The present sequence is that of prawn (Penaeus japonicus) chitinase. This
CC protein can be used in claimed fusion proteins of the invention
CC comprising a translocating moiety and a toxic moiety, where the
CC translocating moiety is a plant protein (e.g. a lectin) capable of acting
CC as a carrier to translocate the toxic moiety across the gut wall of a
CC plant pathogen, and the toxic moiety is an arthropod-derived peptide or
CC protein capable of causing deleterious effects on growth, development,
CC reproduction or mortality in pest insects. Suitable arthropod peptides
CC and proteins include allatostatsins, chitinase, diuretic hormone and their
CC metabolites and analogues. Polynucleotides encoding the fusion protein,
CC vectors, host cells and transgenic plants that are resistant to disease
CC are also provided. The fusion protein is target-specific, and resists
CC degradation in the insect gut. (Updated on 23-OCT-2003 to standardise OS
CC field)
XX
SQ Sequence 527 AA;
XX
Query Match 90.5%; Score 38; DB 6; Length 527;
Best Local Similarity 71.4%; Pred. No. 4.7e+02;
Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 CXCPHP 7
DB 500 CGCEPHP 506

```

```

RESULT 11
ABP72624 ID ABP72624 standard; protein; 527 AA.
XX
AC ABP72624;
XX
DT 23-OCT-2003 (revised)
DT 11-JUN-2003 (first entry)
XX
DE þrawn chitinase.

```

```

RESULT 12
AAU29832 ID AAU29832 standard; protein; 122 AA.
XX
AC AAU29832;
XX
DT 18-DEC-2001 (first entry)
XX
DE Novel human secreted protein #323.
XX
KM Human; vaccination; gene therapy; nutritional supplement;
KM stem cell proliferation; haematopoiesis; nerve tissue regeneration;
KM immune suppression; immune stimulation; anti-inflammatory; leukaemia.
XX
OS Homo sapiens.
XX
PN WO200179449-A2.
XX
PD 25-OCT-2001.

```

```

XX PF 16-APR-2001; 2001WO-US008656.
XX PR 18-APR-2000; 2000US-00552929.
XX PR 26-JUN-2001; 2001US-00770160.
XX PA (HYSE-) HYSEQ INC.
XX PI Tang YT, Liu C, Drmanac RT;
XX DR WPI; 2001-611725/70.
XX PT Nucleic acids encoding a range of human polypeptides, useful in genetic
XX PT vaccination, testing and therapy.
XX PS Claim 20; Page 198; 765pp; English.
XX CC The invention relates to novel human secreted polypeptides. The
XX CC polypeptides and antibodies to the polypeptides are useful for
XX CC determining the presence of or predisposition to a disease associated
XX CC with altered levels of polypeptide. The polypeptides are also useful for
XX CC identifying agents (agonists and antagonists) that bind to them. Cells
XX CC expressing the proteins are useful for identifying a therapeutic agent
XX CC for use in treatment of a pathology related to aberrant expression or
XX CC physiological interactions of the polypeptide. Vectors comprising the
XX CC nucleic acids encoding the polypeptides and cells genetically engineered
XX CC to express them are also useful for producing the proteins. The proteins
XX CC are useful in genetic vaccination, testing and therapy, and can be used
XX CC as nutritional supplements. They may be used to increase stem cell
XX CC proliferation; to regulate haematopoiesis; and in bone, cartilage, tendon
XX CC and/or nerve tissue growth or regeneration; immune suppression and/or
XX CC stimulation; as anti-inflammatory agents; and in treatment of leukaemias.
XX CC AAU95510-AAU33304 represent the amino acid sequences of novel human
XX CC secreted proteins of the invention
XX SQ Sequence 122 AA;

Query Match      88.1%; Score 37; DB 4; Length 122;
Best Local Similarity 71.4%; Pred. No. 2.1e+02;
Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CXCXPHP 7
Db 80 CXCXPHP 86

RESULT 13
ABB89097
ID ABB89097 standard; protein; 170 AA.
XX AC ABB89097;
XX DT 24-MAY-2002 (first entry)
XX DE Human polypeptide SEQ ID NO 1473.
XX KW Cytostatic; immunosuppressive; nootropic; neuroprotective; antiviral;
XX KW antiallergic; hepatotropic; antidiabetic; antiinflammatory; anticancer;
XX KW vulnerrary; anticonvulsant; antibacterial; antifungal; antiparasitic;
XX KW cardiant; gene therapy; cancer; immune disorder; cardiovascular disorder;
XX KW neurological disease; infection; human; secreted protein.
XX OS Homo sapiens.
XX PN WO200190304-A2.
XX PD 29-NOV-2001.
XX PR 18-MAY-2001; 2001WO-US016450.
XX PR 19-MAY-2000; 2000US-0205515P.
XX PA (HUMA-) HUMAN GENOME SCI INC.

```

```

XX PI Birse CE, Rosen CA;
XX DR WPI; 2002-122018/16.
XX DR N-PSDB; ABL89506.
XX PT Novel 1405 isolated polypeptides, useful for diagnosis, treatment and
XX PT prevention of neural, immune system, muscular, reproductive,
XX PT gastrointestinal, pulmonary, cardiovascular, renal and proliferative
XX PT disorders.
XX PS Claim 11; SEQ ID NO 1473; 2081pp + Sequence Listing; English.
XX CC The invention relates to novel genes (ABU89449-ABU90853) and proteins
XX CC (ABB89040-ABB90444) useful for preventing, treating or ameliorating
XX CC medical conditions e.g. by protein or gene therapy. The genes are
XX CC isolated from a range of human tissues disclosed in the specification.
XX CC The nucleic acids, proteins, antibodies and (ant)agonists are useful in
XX CC the diagnosis, treatment and prevention of: (a) cancer, e.g. breast and
XX CC ovarian cancer and other cancers of the adrenal gland, bone, bone marrow,
XX CC breast, gastrointestinal tract, liver, lung, or urogenital; (b) immune
XX CC disorders e.g. Addison's disease, allergies, autoimmune haemolytic
XX CC anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's disease,
XX CC multiple sclerosis, rheumatoid arthritis and ulcerative colitis; (c)
XX CC cardiovascular disorders such as myocardial ischaemias; (d) wound healing
XX CC infectious diseases such as viral, bacterial, fungal and parasitic
XX CC infections. Note: The sequence data for this patent did not form part of
XX CC the printed specification, but was obtained in electronic format directly
XX CC from WPI at ftp.wipo.int/pub/published_pct_sequences
XX SQ Sequence 170 AA;

Query Match      88.1%; Score 37; DB 5; Length 170;
Best Local Similarity 71.4%; Pred. No. 2.7e+02;
Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CXCXPHP 7
Db 114 CXCXPHP 120

RESULT 14
AAU49509
ID AAU49509 standard; protein; 104 AA.
XX AC AAU49509;
XX DT 13-FEB-2002 (first entry)
XX DE Propionibacterium acnes immunogenic protein #10405.
XX KW SAPHO syndrome; synovitis; acne; pustulosis; hyperostosis; osteomyelitis;
XX KW uveitis; endophthalmitis; bone; joint; central nervous system; ELISA;
XX KW inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay;
XX KW dermatological; osteopathic; neuroprotectant.
XX OS Propionibacterium acnes.
XX PN WO200181581-A2.
XX PD 01-NOV-2001.
XX PF 20-APR-2001; 2001WO-US012865.
XX PR 21-APR-2000; 2000US-0199047P.
XX PR 02-JUN-2000; 2000US-0208841P.
XX PR 07-JUL-2000; 2000US-0216747P.
XX PA (CORI-) CORIXA CORP.
XX PI Skeiky YAW, Persing DH, Mitcham JL, Wang SS, Bhatia A;
XX PI L'maisonneuve J, Zhang Y, Jen S, Carter D;

```

```

XX WI, 2001-616774/71.
DR N-PSDB; AAS59545.
XX
XX Propionibacterium acnes polypeptides and nucleic acids useful for
PT vaccinating against and diagnosing infections, especially useful for
PT treating acne vulgaris.
XX
XX Example 1; SEQ ID NO 10704; 1069bp; English.
XX
XX Sequences AAU9105-AAU6807 represent Propionibacterium acnes immunogenic
CC polypeptides. The proteins and their associated DNA sequences are used in
CC the treatment, prevention and diagnosis of medical conditions caused by
CC P. acnes. The disorders include SAPHO syndrome (synovitis, acne,
CC pustulosis, hyperostosis and osteomyelitis), uveitis and endophthalmitis.
CC P. acnes is also involved in infections of bone, joints and the central
CC nervous system, however it is particularly involved in the inflammatory
CC lesions associated with acne vulgaris. A method for detecting the
CC presence or absence of P. acnes in a patient comprises contacting a
CC sample with a binding agent that binds to the proteins of the invention
CC and determining the amount of bound protein in the sample. The
CC polypeptides may be used as antigens in the production of antibodies
CC specific for P. acnes proteins. These antibodies can be used to
CC downregulate expression and activity of P. acnes polypeptides and
CC therefore treat P. acnes infections. The antibodies may also be used as
CC diagnostic agents for determining P. acnes presence, for example, by
CC enzyme linked immunosorbent assay (ELISA). Note: The sequence data for
CC this patent did not form part of the printed specification, but was
CC obtained in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
XX SQ Sequence 104 AA;
XX
XX Query Match 81.0%; Score 34; DB 4; Length 104;
XX Best Local Similarity 57.1%; Pred. No. 5.1e+02;
XX Matches 4; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
XX
XX RESULT 15
XX ABM46028
XX ID ABM46028 standard; protein; 104 AA.
XX
XX AC ABM46028;
XX
XX DT 20-OCT-2003 (first entry)
XX
XX DE Propionibacterium acnes predicted ORF-encoded polypeptide #10704.
XX
XX ACNE vulgaris; antiseborrheic; dermatological; antibacterial;
XX immunostimulant; immune response; vaccine.
XX
XX OS Propionibacterium acnes.
XX
XX PN WO2003033515-A1.
XX
XX PD 24-APR-2003.
XX
XX PF 11-OCT-2002; 2002WO-US032727.
XX
XX PR 15-OCT-2001; 2001US-00978825.
XX
XX PA (CORI-) CORIXA CORP.
XX
XX MITCHELL JL, SKEELY YAW, PERSING DH, BHATIA A, MAISOEUVRE JL,
PI ZHANG Y, WANG S, JEN S, LODES MJ, BENSON DR, JONES R, CARTER D;
PI BARTH B, VALLIEVE-DOUGLASS J;
XX
XX WPI; 2003-381789/36.
XX
XX DR N-PSDB; ACP64474.

```

```

XX
XX New Propionibacterium acnes polypeptides and polynucleotides encoding the
PT polypeptide, useful for diagnosing, preventing or treating acne vulgaris,
PT or for stimulating an immune response specific for a P. acnes protein.
XX
XX Example 1; SEQ ID NO 10704; 1481bp; English.
XX
XX The invention relates to an isolated polynucleotide (ACF64435-ACF64733)
CC encoding a Propionibacterium acnes protein. The invention also relates to
CC polypeptides encoded by the polynucleotides (ABM35624-ABM64536) and to
CC immunogenic fragments of P. acnes polypeptides. The invention
CC additionally encompasses expression vectors and host cells comprising a
CC polynucleotide of the invention; antibodies against polypeptides of the
CC invention; fusion proteins comprising a polypeptide of the invention; a
CC method for stimulating an immune response specific for a P. acnes
CC polypeptide and an isolated T cell population comprising T cells prepared
CC via this method; a vaccine composition (comprising P. acnes polypeptides,
CC polynucleotides, antibodies, fusion proteins, T cell populations, or
CC antigen-presenting cells that express the polypeptide); a method and kit
CC for detecting or determining the presence or absence of P. acnes in a
CC patient; and a method for inhibiting the development of P. acnes in a
CC patient. The P. acnes polypeptides, polynucleotides, antibodies, fusion
CC proteins, T cell populations or antigen-presenting cells that express the
CC polypeptides are useful for diagnosing, preventing or treating acne
CC vulgaris, or for stimulating an immune response specific for a P. acnes
CC protein. The polynucleotides can also be used as probes or primers for
CC nucleic acid hybridisation. The vaccine composition is useful for the
CC stimulation of an immune response against P. acnes, or for treating acne,
CC and the kit is useful for performing a diagnostic assay. The present
CC sequence represents a polypeptide predicted to be encoded by an ORF (open
CC reading frame) contained within the P. acnes polynucleotides of the
CC invention. Note: The sequence data for this patent did not form part of
CC the printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
XX SQ Sequence 104 AA;
XX
XX Query Match 81.0%; Score 34; DB 6; Length 104;
XX Best Local Similarity 57.1%; Pred. No. 5.1e+02;
XX Matches 4; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
XX
XX QY 1 CXKXPHP 7
XX
XX DB 96 CSCSPYP 102
XX
XX Search completed: April 8, 2004, 11:09:54
XX Job time : 46.1417 secs

```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 8, 2004, 11:02:14 ; Search time 12.7361 Seconds
(without alignments)
28.375 Million cell updates/sec

Title: US-09-753-139C-1

Perfect score: 42

Sequence: 1 CKCXPHP 7

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database : Issued Patents AA:*

1: /cgn2_6/prodata/2/1aa/5A_COMB.pep:*
2: /cgn2_6/prodata/2/1aa/5B_COMB.pep:*
3: /cgn2_6/prodata/2/1aa/5A_COMB.pep:*
4: /cgn2_6/prodata/2/1aa/5B_COMB.pep:*
5: /cgn2_6/prodata/2/1aa/PCTUS_COMB.pep:*
6: /cgn2_6/prodata/2/1aa/Backfile1.pep:*

Prod. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	34	81.0	449	4 US-09-252-991A-17669	Sequence 17669, A
2	33	78.6	181	4 US-09-252-991A-30203	Sequence 30203, A
3	33	78.6	592	4 US-08-933-711B-18	Sequence 18, Appl
4	32	76.2	133	4 US-09-489-039A-8115	Sequence 8115, Ap
5	32	76.2	271	1 US-08-152-019A-28	Sequence 28, Appl
6	32	76.2	278	2 US-08-460-309-13	Sequence 13, Appl
7	32	76.2	278	1 US-08-125-077-13	Sequence 13, Appl
8	32	76.2	279	1 US-08-152-019A-29	Sequence 29, Appl
9	32	76.2	501	4 US-09-252-991A-18409	Sequence 18409, A
10	32	76.2	549	3 US-09-245-041-9	Sequence 9, Appl
11	32	76.2	1147	1 US-08-144-121-3	Sequence 3, Appl
12	32	76.2	1147	2 US-08-735-893-3	Sequence 3, Appl
13	32	76.2	1165	1 US-08-144-121-2	Sequence 2, Appl
14	32	76.2	1170	4 US-08-735-893-2	Sequence 2, Appl
15	32	76.2	1170	4 US-09-561-818A-20	Sequence 20, Appl
16	32	76.2	1172	4 US-09-919-172-16	Sequence 12, Appl
17	32	76.2	1260	4 US-09-245-041-2	Sequence 16, Appl
18	32	76.2	1587	3 US-09-845-583A-10	Sequence 10, Appl
19	32	76.2	1587	4 US-09-561-709B-3	Sequence 3, Appl
20	32	76.2	1725	4 US-09-562-702A-20	Sequence 20, Appl
21	32	76.2	1725	4 US-09-561-818A-20	Sequence 20, Appl
22	32	76.2	1765	4 US-09-562-702A-16	Sequence 16, Appl
23	32	76.2	1765	4 US-09-561-818A-16	Sequence 16, Appl
24	32	76.2	1786	4 US-09-562-702A-14	Sequence 14, Appl
25	32	76.2	1786	4 US-09-561-818A-14	Sequence 14, Appl
26	32	76.2	1786	4 US-09-561-818A-14	Sequence 14, Appl
27	32	76.2	1786	4 US-09-561-818A-18	Sequence 18, Appl

28	32	76.2	1786	4 US-09-561-709B-9	Sequence 9, Appl
29	31	73.8	7	1 US-08-179-481-42	Sequence 42, Appl
30	31	73.8	39	1 US-08-096-942-1	Sequence 1, Appl
31	31	73.8	39	5 PCT-US94-08063-1	Sequence 1, Appl
32	31	73.8	52	3 US-09-346-860-3	Sequence 3, Appl
33	31	73.8	52	4 US-09-735-685-3	Sequence 3, Appl
34	31	73.8	65	4 US-09-461-325-457	Sequence 457, App
35	31	73.8	65	4 US-10-013-542-457	Sequence 457, App
36	31	73.8	70	4 US-09-540-236-2438	Sequence 2438, Ap
37	31	73.8	116	3 US-09-081-320-29	Sequence 29, Appl
38	31	73.8	116	4 US-09-574-141A-29	Sequence 29, Appl
39	31	73.8	116	4 US-09-707-780-29	Sequence 29, Appl
40	31	73.8	117	3 US-09-081-320-7	Sequence 7, Appl
41	31	73.8	117	3 US-09-081-320-18	Sequence 7, Appl
42	31	73.8	117	4 US-09-574-141A-7	Sequence 7, Appl
43	31	73.8	117	4 US-09-574-141A-18	Sequence 18, Appl
44	31	73.8	117	4 US-09-574-141A-74	Sequence 74, Appl
45	31	73.8	117	4 US-09-707-780-7	Sequence 7, Appl

ALIGNMENTS

```

RESULT 1
US-09-252-991A-17669
; Sequence 17669, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 17669
; LENGTH: 449
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-17669

Query Match      81.0%; Score 34; DB 4; Length 449;
Best Local Similarity 57.1%; Pred. No. 3.2e+02;
Matches 4; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY      1 CKCXPHP 7
DB      32 CSCSPYP 38

RESULT 2
US-09-252-991A-30203
; Sequence 30203, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 30203
; LENGTH: 181
; TYPE: PRT

```

ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-30203

Query Match 78.6%; Score 33; DB 4; Length 181;
Best Local Similarity 66.7%; Pred. No. 2.1e+02;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CXCXPH 6
DB 91 CACAPH 96

RESULT 3

US-08-933-711B-18
Sequence 18, Application US/08933711B
Patent No. 6514724
GENERAL INFORMATION:
APPLICANT: McMahon, Andrew P.
APPLICANT: Chuang, Pao-Tien
TITLE OF INVENTION: HEDGEHOG INTERACTING PROTEINS AND USES RELATED THERETO
FILE REFERENCE: HUV-024.01
CURRENT APPLICATION NUMBER: US/08/933,711B
CURRENT FILING DATE: 1997-09-19
PRIOR APPLICATION NUMBER: 60/026,155
PRIOR FILING DATE: 1996-09-20
NUMBER OF SEQ ID NOS: 18
SOFTWARE: Patent In Ver. 2.1
SEQ ID NO 18
LENGTH: 592
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Consensus
US-08-933-711B-18

Query Match 78.6%; Score 33; DB 4; Length 592;
Best Local Similarity 66.7%; Pred. No. 5.7e+02;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CXCXPH 6
DB 73 CACSPH 78

RESULT 4

US-09-489-039A-8115
Sequence 8115, Application US/09489039A
Patent No. 6610836
GENERAL INFORMATION:
APPLICANT: Gary Breton et. al
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
FILE REFERENCE: 2709,2004001
CURRENT APPLICATION NUMBER: US/09/489,039A
CURRENT FILING DATE: 2000-01-27
PRIOR APPLICATION NUMBER: US 60/117,747
PRIOR FILING DATE: 1999-01-29
NUMBER OF SEQ ID NOS: 14342
SEQ ID NO 8115
LENGTH: 133
TYPE: PRT
ORGANISM: Klebsiella pneumoniae
US-09-489-039A-8115

Query Match 76.2%; Score 32; DB 4; Length 133;
Best Local Similarity 66.7%; Pred. No. 2.3e+02;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CXCXPH 6
DB 97 CGCSPH 102

RESULT 5

US-08-152-019A-28
Sequence 28, Application US/08152019A
Patent No. 556331
GENERAL INFORMATION:
APPLICANT: Tessier-Lavigne, Marc
APPLICANT: Serafini, Tito
APPLICANT: Kennedy, Timothy
APPLICANT: Placzek, Marysia
APPLICANT: Jessell, Thomas
TITLE OF INVENTION: NEURAL AXON OUTGROWTH MODULATORS
NUMBER OF SEQUENCES: 46
CORRESPONDENCE ADDRESS:
ADDRESSEE: FLEHR, HOBBACH, TEST, ALBRITTON & HERBERT
STREET: 4 Embarcadero Center, Suite 3400
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111-4187
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/152,019A
FILING DATE: 12-NOV-1993
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: Osman, Richard Aron
REGISTRATION NUMBER: 36,627
REFERENCE/DOCKET NUMBER: A-59012/RAO
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 781-1989
TELEX: 910 277299 FHT UR
INFORMATION FOR SEQ ID NO: 28:
SEQUENCE CHARACTERISTICS:
LENGTH: 271 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-152-019A-28

Query Match 76.2%; Score 32; DB 1; Length 271;
Best Local Similarity 66.7%; Pred. No. 4.2e+02;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CXCXPH 6
DB 260 CSCDPH 265

RESULT 6

US-08-460-309-13
Sequence 13, Application US/08460309
Patent No. 5837496
GENERAL INFORMATION:
APPLICANT: Engvall, Eva
APPLICANT: Leivo, Ilmo
TITLE OF INVENTION: Nucleic Acids Encoding Merosin, Merosin
NUMBER OF SEQUENCES: 23
CORRESPONDENCE ADDRESS:
STREET: 4370 La Jolla Village Drive, Suite 700
CITY: San Diego
STATE: California
COUNTRY: USA
ZIP: 92122

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/460,309
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/125,077
FILING DATE: 22-SEP-1993
APPLICATION NUMBER: US PCT/US 94/10730
FILING DATE: 21-SEP-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/472,319
FILING DATE: 30-JAN-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/919,951
FILING DATE: 27-JUL-1992
ATTORNEY/AGENT INFORMATION:
NAME: Campbell, Cathryn A.
REGISTRATION NUMBER: 31,815
REFERENCE/DOCKET NUMBER: P-LA 9721
TELEPHONE: (619) 535-9001
TELEFAX: (619) 535-8949
INFORMATION FOR SEQ ID NO: 13:
SEQUENCE CHARACTERISTICS:
LENGTH: 278 amino acids
TYPE: amino acid
TOPOLOGY: linear
US-08-460-309-13

Query Match 76.2%; Score 32; DB 2; Length 278;
Best Local Similarity 66.7%; Pred. No. 4.3e+02;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CXCRPH 6
DB 259 CSCRPH 264

RESULT 7
US-08-125-077-13
Sequence 13, Application US/08125077
Patent No. 5872231
Patent No. 5872231 5840863
GENERAL INFORMATION:
APPLICANT: Engvall, Eva
APPLICANT: Leivo, Ilmo
TITLE OF INVENTION: Nucleic Acids Encoding Merosin, Merosin
TITLE OF INVENTION: Fragments and Uses Thereof
NUMBER OF SEQUENCES: 23
CORRESPONDENCE ADDRESS:
ADDRESSEE: Campbell and Flores
STREET: 4370 La Jolla Village Drive, Suite 700
CITY: San Diego
STATE: California
COUNTRY: USA
ZIP: 92122
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/125,077
FILING DATE: 22-SEP-1993
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US PCT/US 94/10730
FILING DATE: 21-SEP-1994

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/472,319
FILING DATE: 30-JAN-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/919,951
FILING DATE: 27-JUL-1992
ATTORNEY/AGENT INFORMATION:
NAME: Campbell, Cathryn A.
REGISTRATION NUMBER: 31,815
REFERENCE/DOCKET NUMBER: P-LA 9721
TELEPHONE: (619) 535-9001
TELEFAX: (619) 535-8949
INFORMATION FOR SEQ ID NO: 13:
SEQUENCE CHARACTERISTICS:
LENGTH: 278 amino acids
TYPE: amino acid
TOPOLOGY: linear
US-08-125-077-13

Query Match 76.2%; Score 32; DB 2; Length 278;
Best Local Similarity 66.7%; Pred. No. 4.3e+02;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CXCRPH 6
DB 259 CSCRPH 264

RESULT 8
US-08-152-019A-29
Sequence 29, Application US/08152019A
Patent No. 5565351
GENERAL INFORMATION:
APPLICANT: Tessier-Lavigne, Marc
APPLICANT: Serafini, Tito
APPLICANT: Kennedy, Timothy
APPLICANT: Placzek, Marysia
APPLICANT: Jessell, Thomas
APPLICANT: Dodd, Jane
TITLE OF INVENTION: NEURAL AXON OUTGROWTH MODULATORS
NUMBER OF SEQUENCES: 46
CORRESPONDENCE ADDRESS:
ADDRESSEE: FLEHR, HOBBACH, TEST, ALBRITTON & HERBERT
STREET: 4 Embarcadero Center, Suite 3400
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111-4187
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/152,019A
FILING DATE: 12-NOV-1993
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: Osman, Richard Aron
REGISTRATION NUMBER: 36,627
REFERENCE/DOCKET NUMBER: A-59012/BAO
TELEPHONE: (415) 781-1989
TELEFAX: (415) 398-3249
TELEX: 910 277299 FHR UR
INFORMATION FOR SEQ ID NO: 29:
SEQUENCE CHARACTERISTICS:
LENGTH: 279 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide

US-08-152-019A-29

Query Match 76.2%; Score 32; DB 1; Length 279;
Best Local Similarity 66.7%; Pred. No. 4.3e+02;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 CXCPH 6
Db 260 CSCRP 265

RESULT 9

US-09-252-991A-18409
; Sequence 18409, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107136.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 18409
; LENGTH: 501
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-18409

Query Match 76.2%; Score 32; DB 4; Length 501;
Best Local Similarity 57.1%; Pred. No. 7e+02;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 CXCPH 7
Db 132 CTCRPR 138

RESULT 10

US-09-245-041-9
; Sequence 9, Application US/09245041
; Patent No. 6274339
; GENERAL INFORMATION:
; APPLICANT: Moore, K.
; APPLICANT: Nagle, D.
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE DIAGNOSIS AND TREATMENT
; FILE REFERENCE: 7853-136
; CURRENT APPLICATION NUMBER: US/09/245,041
; CURRENT FILING DATE: 1999-02-05
; EARLIER APPLICATION NUMBER: 60/093,630
; EARLIER FILING DATE: 1998-07-21
; EARLIER APPLICATION NUMBER: 60/104,978
; EARLIER FILING DATE: 1998-10-20
; NUMBER OF SEQ ID NOS: 131
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 9
; LENGTH: 549
; TYPE: PRT
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: SITE
; LOCATION: all Xaa positions
; OTHER INFORMATION: Xaa=unknown amino acid
US-09-245-041-9

Query Match 76.2%; Score 32; DB 3; Length 549;
Best Local Similarity 66.7%; Pred. No. 7.6e+02;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 CXCPH 6
Db 138 CSCRP 143

RESULT 11
US-08-144-121-3
; Sequence 3, Application US/08144121
; Patent No. 5610031
; GENERAL INFORMATION:
; APPLICANT: Burgeson, Robert E.
; APPLICANT: Wagnon, David W.
; TITLE OF INVENTION: BIK CHAIN OF LAMININ AND METHODS OF USE
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD
; STREET: 60 State Street, suite 510
; CITY: BOSTON
; STATE: Massachusetts
; COUNTRY: United States
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/144,121
; FILING DATE: 27-OCT-1993
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Myers, Paul L.
; REGISTRATION NUMBER: 35,965
; REFERENCE/DOCKET NUMBER: (MGP-0780.0) MGP-021
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 227-7400
; TELEFAX: (617) 227-5941
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1147 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FRAGMENT TYPE: internal
; FEATURE:
; NAME/KEY: Domain
; LOCATION: 1..231
; FEATURE:
; NAME/KEY: Domain
; LOCATION: 232..411
; FEATURE:
; NAME/KEY: Domain
; LOCATION: 412..765
; FEATURE:
; NAME/KEY: Domain
; LOCATION: 766..1147
US-08-144-121-3

Query Match 76.2%; Score 32; DB 1; Length 1147;
Best Local Similarity 66.7%; Pred. No. 1.4e+03;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 CXCPH 6
Db 461 CACDPH 466

RESULT 12
US-08-735-893-3
; Sequence 3, Application US/08735893
; Patent No. 5914317
; GENERAL INFORMATION:

APPLICANT: Burgeson, Robert E.
TITLE OF INVENTION: B1K CHAIN OF LAMININ AND METHODS OF USE
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:
ADDRESSEE: LAHIVE & COCKFIELD
STREET: 60 State Street, suite 510
CITY: BOSTON
STATE: Massachusetts
COUNTRY: United States
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/735,893
FILING DATE: 18-OCT-1996
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/144,121
FILING DATE: 27-OCT-1993
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Myers, Paul L.
REGISTRATION NUMBER: 35,965
REFERENCE/DOCKET NUMBER: (MGP-0780.1) MGP-021DV
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 227-7400
TELEFAX: (617) 227-5941
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 1147 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
FRAGMENT TYPE: internal
FEATURE:
NAME/KEY: Domain
LOCATION: 1..231
FEATURE:
NAME/KEY: Domain
LOCATION: 232..411
FEATURE:
NAME/KEY: Domain
LOCATION: 412..765
FEATURE:
NAME/KEY: Domain
LOCATION: 766..1147
US-08-735-893-3
Query Match 76.2%; Score 32; DB 2; Length 1147;
Best Local Similarity 66.7%; Pred. No. 1.4e+03;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 CKCXPB 6
DB 461 CACDPB 466
RESULT 13
US-08-144-121-2
Sequence 2, Application US/08144121
Patent No. 5610031
GENERAL INFORMATION:
APPLICANT: Burgeson, Robert E.
TITLE OF INVENTION: B1K CHAIN OF LAMININ AND METHODS OF USE
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:
ADDRESSEE: LAHIVE & COCKFIELD
STREET: 60 State Street, suite 510

CITY: BOSTON
STATE: Massachusetts
COUNTRY: United States
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/144,121
FILING DATE: 27-OCT-1993
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Myers, Paul L.
REGISTRATION NUMBER: 35,965
REFERENCE/DOCKET NUMBER: (MGP-0780.0) MGP-021
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 227-7400
TELEFAX: (617) 227-5941
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 1165 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-144-121-2
Query Match 76.2%; Score 32; DB 1; Length 1165;
Best Local Similarity 66.7%; Pred. No. 1.4e+03;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 CKCXPB 6
DB 478 CACDPB 483
RESULT 14
US-08-735-893-2
Sequence 2, Application US/08735893
Patent No. 5914317
GENERAL INFORMATION:
APPLICANT: Burgeson, Robert E.
TITLE OF INVENTION: B1K CHAIN OF LAMININ AND METHODS OF USE
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:
ADDRESSEE: LAHIVE & COCKFIELD
STREET: 60 State Street, suite 510
CITY: BOSTON
STATE: Massachusetts
COUNTRY: United States
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/735,893
FILING DATE: 18-OCT-1996
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/144,121
FILING DATE: 27-OCT-1993
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Myers, Paul L.
REGISTRATION NUMBER: 35,965
REFERENCE/DOCKET NUMBER: (MGP-0780.1) MGP-021DV
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 227-7400
TELEFAX: (617) 227-5941

; INFORMATION FOR SEQ ID NO: 2;
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1165 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-735-893-2

Query Match 76.2%; Score 32; DB 2; Length 1165;
Best Local Similarity 66.7%; Pred. No. 1.4e+03;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CXCPH 6
Db 478 CACDPH 483

RESULT 15
US-09-561-709B-12
; Sequence 12, Application US/09561709B
; Patent No. 6682911
; GENERAL INFORMATION:
; APPLICANT: Burgeson, Robert
; APPLICANT: Champlaud, Marie-France
; APPLICANT: Olson, Pamela
; APPLICANT: Koch, Manuel
; APPLICANT: Brunken, William
; TITLE OF INVENTION: LAMININS AND USES THEREOF
; FILE REFERENCE: 10287-06001
; CURRENT APPLICATION NUMBER: US/09/561,709B
; CURRENT FILING DATE: 2000-05-01
; PRIOR APPLICATION NUMBER: US 09/168,949
; PRIOR FILING DATE: 1998-10-09
; PRIOR APPLICATION NUMBER: US 60/061,609
; PRIOR FILING DATE: 1997-10-10
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12
; LENGTH: 1170
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-561-709B-12

Query Match 76.2%; Score 32; DB 4; Length 1170;
Best Local Similarity 66.7%; Pred. No. 1.4e+03;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CXCPH 6
Db 479 CACDPH 484

Search completed: April 8, 2004, 11:20:27
Job time : 13.7361 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using SW model

Run on: April 8, 2004, 11:01:14 ; Search time 9.52778 Seconds
(without alignments)
70.671 Million cell updates/sec

Title: US-09-753-139C-1

Perfect score: 42
Sequence: 1 CKCXPDP 7

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :
1: p1r1:*
2: p1r2:*
3: p1r3:*
4: p1r4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	33	78.6	158	2 T04875	hypothetical prote
2	33	78.6	215	2 T19056	hypothetical prote
3	32	76.2	103	2 H72753	hypothetical prote
4	32	76.2	119	2 E89101	protein F25E5.9 [l
5	32	76.2	1170	2 A53612	laminin Blk chain
6	32	76.2	1376	1 VGIHJ2	E2 glycoprotein pr
7	32	76.2	1376	1 J01534	E2 glycoprotein pr
8	32	76.2	1786	1 MMHUB1	laminin beta-1 cha
9	32	76.2	1786	1 MMSB1	laminin beta-1 cha
10	31	73.8	39	1 A48523	margatoxin [valida
11	31	73.8	68	2 JV0044	hypothetical 7.5K
12	31	73.8	87	2 A35666	transcription acti
13	31	73.8	119	1 CUPSAM	amicyanin precuro
14	31	73.8	131	1 S12972	amicyanin - Paraco
15	31	73.8	132	1 A23706	amicyanin - Thioba
16	31	73.8	145	2 T36527	hypothetical prote
17	31	73.8	149	2 F87579	hypothetical prote
18	31	73.8	211	2 T04119	probable berine/ch
19	31	73.8	212	2 S53257	e anti-gen precuro
20	31	73.8	235	2 JCA603	conerved hypotet
21	31	73.8	253	2 S40181	phosphoprotein pho
22	31	73.8	253	2 T25768	phosphoprotein pho
23	31	73.8	282	2 T15640	hypothetical prote
24	31	73.8	313	1 FOVDA	hypothetical prote
25	31	73.8	317	1 MMKR3S	gag polypeptin
26	31	73.8	317	1 MMKR3B	34K protein - simi
27	31	73.8	317	1 MMKR5U	34K protein - boyl
28	31	73.8	317	2 S30581	gp33 protein - hum
29	31	73.8	317	2 S49005	non-structural pro

30	31	73.8	343	2 T29547	hypothetical prote
31	31	73.8	352	2 T24369	hypothetical prote
32	31	73.8	355	2 T15797	hypothetical prote
33	31	73.8	448	2 T15188	hypothetical prote
34	31	73.8	454	2 T03130	probable tyrosine
35	31	73.8	458	2 S12444	hypothetical prote
36	31	73.8	459	2 F71257	hypothetical prote
37	31	73.8	486	1 JC7241	phosphoprotein pho
38	31	73.8	502	1 UC1283	phosphoprotein pho
39	31	73.8	506	2 A54190	cerebroside-sulfat
40	31	73.8	509	2 OKHULK	protein-tyrosine k
41	31	73.8	513	2 A38193	phosphoprotein pho
42	31	73.8	514	1 A36222	phosphoprotein pho
43	31	73.8	515	2 JT0976	phosphoprotein pho
44	31	73.8	518	2 A40942	phosphoprotein pho
45	31	73.8	521	1 S35067	phosphoprotein pho

ALIGNMENTS

RESULT 1
T04875
hypothetical protein F18P4.10 - Arabidopsis thaliana (fragment)
C/Species: Arabidopsis thaliana (mouse-ear cress)
C/Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #ext_change 11-Jun-1999
C/Accession: T04875
R/Bevan, M.; Murphy, G.; Ridley, P.; Hudson, S.; Bancroft, I.; Mewes, H.W.; Mayer, K.F.X
submitted to the Protein Sequence Database, February 1998
A/Reference number: Z15388
A/Accession: T04875
A/Molecule type: DNA
A/Residues: 1-158 <BEV>
A/Cross-references: EMBL:AL021637
A/Experimental source: cultivar Columbia; BAC clone F18P4
C/Genetics:
A/Map position: 4
A/Intons: 18/1; 72/1; 121/3
A/Note: F18P4.10

Query Match
Best Local Similarity 78.6%; Score 33; DB 2; Length 158;
Matches 4; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 CKCXPDP 7
DB 70 CTCTPDP 76

RESULT 2
T19056
hypothetical protein C07E3.6 - Caenorhabditis elegans
C/Species: Caenorhabditis elegans
C/Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #ext_change 09-Jun-2000
C/Accession: T19056
R/Matthews, P.
submitted to the EMBL Data Library, June 1995
A/Reference number: Z19066
A/Accession: T19056
A/Molecule type: DNA
A/Residues: 1-215 <WIL>
A/Cross-references: EMBL:Z49908; PIRN:CAA90100.1; GSPDB:GN00020; CESP:C07E3.6
A/Experimental source: clone C07E3
C/Genetics:
A/Map position: 2
A/Intons: 28/3; 136/3; 162/3; 190/3
C/Superfamily: Caenorhabditis elegans hypothetical protein C07E3.6
Query Match
Best Local Similarity 78.6%; Score 33; DB 2; Length 215;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 CXCPH 6
Db 49 CSCTPH 54

RESULT 3

H72753
hypothetical protein APE0021 - Aeropyrum pernix (strain K1)
C:Species: Aeropyrum pernix
C:Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 20-Aug-1999
C:Accession: H72753
R:Kawarayashi, Y.; Hino, Y.; Horikawa, H.; Yamazaki, S.; Haikawa, Y.; Jin-no, K.; Takahawa, H.; Takamiya, M.; Maeda, S.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; K DNA Res. 6, 83-101, 1999
A>Title: Complete genome sequence of an aerobic hyper-thermophilic Crenarchaeon, Aeropyrum pernix
A:Reference number: A12450; MUID:99310339; PMID:10382966
A:Accession: H72753
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-103 <KAW>
A:Cross-references: DDBJ:AP000058; NID:G5103388; PIDN:BAH78930.1; PID:dl042706; PID:G510
A:Experimental source: strain K1
C:Genetics:
A:Gene: APE0021

Query Match 76.2%; Score 32; DB 2; Length 103;
Best Local Similarity 57.1%; Pred. No. 1e+02;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 CXCPH 7
Db 13 CSCTPH 19

RESULT 4

E89101
protein P25B5.9 [imported] - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 10-May-2001
C:Accession: E89101
R:anonymus, The C. elegans Sequencing Consortium.
Science 282, 2012-2018, 1998
A>Title: Genome sequence of the nematode C. elegans: a platform for investigating biology
A:Reference number: A75000; MUID:99069613; PMID:9851916
A>Note: see web site genome.wustl.edu/gsc/C_elegans/ and www.sanger.ac.uk/Projects/C_eleg
A>Note: published errata appeared in Science 283, 35, 1999; Science 283, 2103, 1999; and
A:Accession: E89101
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-119 <STO>
A:Cross-references: GB:chr_V; PIDN:AAC27339.1; PID:G3335246; GSPDB:GN00023; CESP:P25B5.9
C:Genetics:
A:Gene: P25B5.9
A:Map position: 5

Query Match 76.2%; Score 32; DB 2; Length 119;
Best Local Similarity 66.7%; Pred. No. 1.2e+02;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 CXCPH 6
Db 71 CMCAPH 76

RESULT 5

A53612
laminin B1k chain precursor - human
N:Alternate names: kalinin B1 chain, nicein B1 chain
C:Species: Homo sapiens (man)
C:Date: 07-Oct-1994 #sequence_revision 07-Oct-1994 #text_change 05-Nov-1999
C:Accession: A53612
R:Gerecke, D.R.; Wegman, D.W.; Champilaud, M.F.; Burgeson, R.E.

J. Biol. Chem. 269, 11073-11080, 1994
A>Title: The complete primary structure for a novel laminin chain, the laminin B1k chain
A:Reference number: A53612; MUID:94209274; PMID:7512558
A:Accession: A53612

A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-1170 <GER>
A:Cross-references: GB:I25541; NID:G510702; PIDN:AAA61834.1; PID:G510703
A>Note: authors translated the codon CGA for residue 124 as Gln, GAG for residue 439 as .
C:Superfamily: laminin-type EGF-like homology

C:Keywords: glycoprotein
F:1-17/Domain: signal sequence #status predicted <SIG>
F:18-1170/Product: laminin B1k chain #status predicted <MAT>
F:250-312/Domain: laminin-type EGF-like homology <LEG1>
F:378-427/Domain: laminin-type EGF-like homology <LEG>
F:430-476/Domain: laminin-type EGF-like homology <EG11>
F:532-576/Domain: laminin-type EGF-like homology <LEG2>

Query Match 76.2%; Score 32; DB 2; Length 1170;
Best Local Similarity 66.7%; Pred. No. 6.8e+02;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 CXCPH 6
Db 479 CACDPH 484

RESULT 6

VG1H72
E2 glycoprotein precursor - murine hepatitis virus (strain wild type MHV-4)
N:Alternate names: peplomer glycoprotein; spike glycoprotein
N:Contains: 90A glycoprotein; 90B glycoprotein
C:Species: murine hepatitis virus, MHV
C:Date: 31-Mar-1991 #sequence_revision 31-Mar-1991 #text_change 16-Jul-1999
C:Accession: A33748
R:Parker, S.E.; Gallagher, T.M.; Buchmeier, M.J.
Virology 173, 664-673, 1989
A>Title: Sequence analysis reveals extensive polymorphism and evidence of deletions with
A:Reference number: A33748; MUID:90085815; PMID:2556846

A:Accession: A33748
A:Molecule type: genomic RNA
A:Residues: 1-1376 <PBR>
A:Cross-references: GB:M32789; NID:G331846; PIDN:AAA46456.1; PID:G331847
C:Superfamily: coronavirus E2 glycoprotein
C:Keywords: glycoprotein; transmembrane protein

F:1-14/Domain: signal sequence #status predicted <SIG>
F:15-1376/Product: E2 glycoprotein #status predicted <EG>
F:15-769/Product: 90B glycoprotein #status predicted <EGB>
F:770-1376/Product: 90A glycoprotein #status predicted <EGA>
F:1321-1338/Domain: transmembrane #status predicted <TMN>
F:31.60,134,192,357,435,442,582,677,709,717,740,789,806,896,945,1178,1232,1242,1261,1277

Query Match 76.2%; Score 32; DB 1; Length 1376;
Best Local Similarity 57.1%; Pred. No. 7.7e+02;
Matches 4; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 CXCPH 7
Db 532 CTCKNP 538

RESULT 7

J01534
E2 glycoprotein precursor - murine hepatitis virus (strain JHM cl-2)
N:Alternate names: peplomer glycoprotein; spike glycoprotein
N:Contains: 90A glycoprotein; 90B glycoprotein
C:Species: murine hepatitis virus, MHV
C:Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 07-May-1999
C:Accession: J01534
R:Rauch, F.; Ikeda, T.; Shida, H.
J. Gen. Virol. 73, 1065-1072, 1992
A>Title: Molecular cloning and expression of a spike protein of neurovirulent murine cor
A:Reference number: J01534; MUID:92268864; PMID:1316938

A:Accession: J01534
 A:Molecule type: mRNA
 A:Residues: 1-1376 <TAG>
 A:Cross-references: DDBJ:D10255
 A>Note: the authors translated the codon TTT for residue 8 as Leu, GCG for residue 14 as Arg
 C:Superfamily: coronavirus E2 glycoprotein
 C:Keywords: glycoprotein; transmembrane protein
 F:1-14/Domain: signal sequence #status predicted <SIG>
 F:15-1376/Product: E2 glycoprotein #status predicted <E2G>
 F:15-769/Product: 908 glycoprotein #status predicted <EGB>
 F:770-1376/Product: 90A glycoprotein #status predicted <EGA>
 F:138-1339/Domain: transmembrane #status predicted <TM>
 F:31,60,134,192,357,435,677,709,717,789,806,945,1232,1242,1261,1277,1298/Binding site: C

Query Match 76.2%; Score 32; DB 1; Length 1376;
 Best Local Similarity 57.1%; Pred. No. 7, 7e+02;
 Matches 4; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 CKCXP 7
 Db 532 CKCXP 538

RESULT 8
 MMSB1
 laminin beta-1 chain precursor - human
 N:Alternate names: laminin chain B1
 C:Species: Homo sapiens (man)
 C:Date: 30-Jun-1991 #sequence revision 30-Jun-1991 #text change 19-Jan-2001
 C:Accession: S13547; A28483; A26994; S23566
 R:Vuolteenaho, R.; Chow, L.T.; Tytgvaasen, K.
 J. Biol. Chem. 265, 15611-15616, 1990
 A>Title: Structure of the human laminin B1 chain gene.
 A:Reference number: S13547; MUID:90368768; PMID:1975589
 A:Accession: S13547
 A:Status: nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-1786 <VUD>
 A:Cross-references: GB:M61951; GB:J02778; NID:G186911; PIDN:AA59486.1; PID:G186913
 A>Note: the nucleotide sequence was submitted to Genbank, February 1991
 R:Piikarainen, T.; Eddy, R.; Fukushima, Y.; Byers, M.; Shows, T.; Pihlajaniemi, T.; Sara J. Biol. Chem. 262, 10454-10462, 1987
 A>Title: Human laminin B1 chain. A multidomain protein with gene (LAMB1) locus in the q2
 A:Reference number: A28483; MUID:87280097; PMID:3611077
 A:Accession: A28483
 A:Molecule type: mRNA
 A:Residues: 1-1786 <PIK>
 A:Cross-references: GB:M61951; GB:J02778; NID:G186911; PIDN:AA59486.1; PID:G186913
 R:Jaye, M.; Modi, W.S.; Ricca, G.A.; Mudd, R.; Chiu, I.M.; O'Brien, S.J.; Drohan, W.N. Am. J. Hum. Genet. 41, 605-615, 1987
 A>Title: Isolation of a cDNA clone for the human laminin-B1 chain and its gene localizat
 A:Reference number: A26994; MUID:88021029; PMID:3661559
 A:Accession: A26994
 A:Molecule type: mRNA
 A:Residues: 1276-1469, 'V', 1471-1695, 'G', 1697-1709 <JAY>
 R:Vuolteenaho, R.; Kallunki, T.; Chow, L.; Ikonen, J.; Piikarainen, T.; Tytgvaasen, K. In Extracellular Matrix Genes, Sandell L.J., and Boyd C.D., eds., pp. 175-193, Academic P
 A>Title: Genes for the human laminin B1 and B2 chains.
 A:Reference number: S23566
 A:Accession: S23566
 A:Molecule type: DNA
 A:Residues: 762-1786 <VU2>
 A>Note: mRNA was also sequenced
 C:Genetics:
 A:Gene: GDB:LAMB1
 A:Cross-references: GDB:119357; OMIM:150240
 A:Map position: 7q31.1-7q31.3
 A:Introns: 13/1, 71/3, 117/1, 141/3, 204/3, 226/1, 233/3, 334/1, 397/1, 457/1, 494/3, 52
 64/3, 1513/1, 1582/2, 1629/3, 1688/3, 1742/1
 C:Complex: Laminins are trimers of an alpha-type, a beta-type, and a gamma-type laminin
 C:Function:
 A:Description: Interact with cells and with other basement membrane proteins to promote

C:Superfamily: laminin beta-1 chain, laminin-type EGF-like homology
 C:Keywords: basement membrane; calcium binding; cell binding; coiled coil; extracellular
 F:1-21/Domain: signal sequence #status predicted <SIG>
 F:122-1786/Product: laminin beta-1 chain #status predicted <MAT>
 F:271-548/Domain: VI <DOM5>
 F:271-548/Domain: VI <DOM5>
 F:335-335/Domain: laminin-type EGF-like homology <LEB01>
 F:335-335/Domain: laminin-type EGF-like homology <LEB02>
 F:338-455/Domain: laminin-type EGF-like homology <LEB03>
 F:458-507/Domain: laminin-type EGF-like homology <LEB04>
 F:463-468/Region: cell adhesion #status predicted
 F:510-540/Domain: laminin-type EGF-like homology #status atypical <LEB05>
 F:549-774/Domain: IV <DOM4>
 F:662-668/Region: cell adhesion #status predicted
 F:773-818/Domain: laminin-type EGF-like homology <LEB06>
 F:773-818/Domain: III <DOM3>
 F:821-864/Domain: laminin-type EGF-like homology <LEB07>
 F:867-914/Domain: laminin-type EGF-like homology <LEB08>
 F:917-973/Domain: laminin-type EGF-like homology <LEB09>
 F:923-927/Region: cell adhesion #status predicted
 F:950-954/Region: cell adhesion #status predicted
 F:976-1025/Domain: laminin-type EGF-like homology <LEB10>
 F:1026-1081/Domain: laminin-type EGF-like homology <LEB11>
 F:1084-1129/Domain: laminin-type EGF-like homology <LEB12>
 F:1133-1176/Domain: laminin-type EGF-like homology <LEB13>
 F:1179-1397/Domain: II <DOM2>
 F:1179-1397/Region: heptad repeats
 F:1398-1430/Domain: alpha <ALP>
 F:1431-1786/Domain: I <DOM1>
 F:1431-1786/Region: heptad repeats
 F:330-35/Disulfide bonds: #status predicted
 F:120,356,519,677,965,1041,1195,1279,1336,1343,1487,1542,1643/Binding site: carbohydrate
 F:1179,1182,1785/Disulfide bonds: Interchain #status predicted

Query Match 76.2%; Score 32; DB 1; Length 1786;
 Best Local Similarity 66.7%; Pred. No. 9, 4e+02;
 Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CKCXP 6
 Db 539 CKCXP 534

RESULT 9
 MMSB1
 laminin beta-1 chain precursor - mouse
 N:Alternate names: laminin chain B1
 C:Species: Mus musculus (house mouse)
 C:Date: 28-Feb-1986 #sequence revision 30-Jun-1991 #text change 10-Dec-1999
 C:Accession: A26413; S02679; S05326; S14877; A02871; S02036; S13543
 R:Sasaki, M.; Kato, S.; Kohno, K.; Martin, G.R.; Yamada, Y. Proc. Natl. Acad. Sci. U.S.A. 84, 935-939, 1987
 A>Title: Sequence of the cDNA encoding the laminin B1 chain reveals a multidomain protei
 A:Reference number: A26413; MUID:87147212; PMID:3493487
 A:Accession: A26413
 A:Molecule type: mRNA
 A:Residues: 1-1786 <SAS>
 A:Cross-references: EMBL:M15525; NID:G198700
 A>Note: translation in GenBank has additional 48 residues at the amino end
 R:Fujiwara, S.; Shinkai, H.; Deutzmann, R.; Paulsson, M.; Timpl, R. Biochem. J. 252, 453-461, 1988
 A>Title: Structure and distribution of N-linked oligosaccharide chains on various domain
 A:Reference number: S02679; MUID:88326259; PMID:2458101
 A:Accession: S02679
 A:Molecule type: protein
 A:Residues: 28-42,932-946 <FUJ>
 R:Harcl, L.; Oberbauer, I.; Deutzmann, R. Eur. J. Biochem. 173, 629-635, 1988
 A>Title: The N termini of laminin A chain is homologous to the B chains.
 A:Accession: S05326
 A:Accession: S05326
 A:Molecule type: protein
 A:Residues: 457-466,854-868,932-946 <HAR>

R.Mann, K.; Deutzmann, R.; Timp1, R.
 Eur. J. Biochem. 178, 71-80, 1988
 A>Title: Characterization of proteolytic fragments of the laminin-nidogen complex and th
 A:Reference number: S08895; PMID:89078415; PMID:2462498
 A:Accession: S14877
 A:Molecule type: protein
 A:Residues: 590-620 <MAN>
 R:Barlow, D.P.; Green, N.M.; Kurkinen, M.; Hogan, B.L.M.
 EMBO J. 3, 2355-2362, 1984
 A>Title: Sequencing of laminin B chain cDNAs reveals C-terminal regions of coiled-coil a
 A:Reference number: A02870; PMID:85051302; PMID:6209134
 A:Accession: A02871
 A:Molecule type: mRNA
 A:Residues: 1292-1530, 'MENP', 1535-1691, 'C', 1693-1748, 'N', 1750-1786 <BAR>
 A:Cross-references: EMBL:X05212; NID:952861; PID:CA28839.1; PID:9809042
 R:Deutzmann, R.; Heber, U.; Schmetz, K.A.; Oberbeumer, I.; Hartl, L.
 Eur. J. Biochem. 177, 35-45, 1988
 A>Title: Structural study of long arm fragments of laminin. Evidence for repetitive C-te
 A:Reference number: S01790; PMID:89030693; PMID:3181157
 A:Accession: S02036
 A:Molecule type: protein
 A:Residues: 1561-1587 <DEU>
 R:Paulsson, M.; Deutzmann, R.; Timp1, R.; Dalzoppo, D.; Odermatt, E.; Engel, J.
 EMBO J. 4, 309-316, 1985
 A>Title: Evidence for coiled-coil alpha-helical regions in the long arm of laminin.
 A:Reference number: S13543; PMID:85257455; PMID:3848400
 A:Accession: S13543
 A:Molecule type: protein
 A:Residues: 1700-1748, 'N', 1750-1759 <PAU>
 C:Genetics:
 A:Gene: Lamb-1
 A:Map position: 12
 C:Complex: Laminins are trimers of an alpha-type, a beta-type, and a gamma-type laminin
 C:Function:
 A:Description: interact with cells and with other basement membrane proteins to promote
 C:Superfamily: laminin beta-1 chain; laminin-type EGF-like homology
 C:Keywords: basement membrane; calcium binding; cell binding; coiled coil; extracellular
 F:1-21/Domain: signal sequence #status predicted <SIG>
 F:12-1786/Product: laminin beta-1 chain #status predicted <MAT>
 F:12-270/Domain: VI <DOM6>
 F:1271-540/Domain: V <DOM5>
 F:1271-332/Domain: laminin-type EGF-like homology <LE01>
 F:133-355/Domain: laminin-type EGF-like homology <LE02>
 F:1398-455/Domain: laminin-type EGF-like homology <LE03>
 F:1458-507/Domain: laminin-type EGF-like homology <LE04>
 F:510-540/Domain: laminin-type EGF-like homology #status atypical <LE05>
 F:541-772/Domain: IV <DOM4>
 F:773-1182/Domain: III <DOM3>
 F:773-818/Domain: laminin-type EGF-like homology <LE06>
 F:1821-864/Domain: laminin-type EGF-like homology <LE07>
 F:1867-914/Domain: laminin-type EGF-like homology <LE08>
 F:1917-973/Domain: laminin-type EGF-like homology <LE09>
 F:1976-1025/Domain: laminin-type EGF-like homology <LE10>
 F:1028-1081/Domain: laminin-type EGF-like homology <LE11>
 F:1084-1129/Domain: laminin-type EGF-like homology <LE12>
 F:11132-1176/Domain: laminin-type EGF-like homology <LE13>
 F:1183-1397/Domain: II <DOM2>
 F:1183-1397/Region: heptad repeats
 F:11398-1430/Domain: alpha <ALP>
 F:1431-1786/Region: heptad repeats
 F:1431-1786/Domain: I <DOM1>
 F:12/Modified site: pyrolydine carboxylic acid (Gln) (in mature form) #status predicted
 F:30-35/Disulfide bonds: #status predicted
 F:1120,356,519,677,1041,1195,1279,1336,1343,1487,1533,1542,1643/Binding site: carbohydrate
 F:1179,1182,1785/Disulfide bonds: interchain #status predicted

Query Match 76.2%; Score 32; DB 1; Length 1786;
 Best Local Similarity 66.7%; Pred. No. 9.4e+02;
 Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CXKXP 6
 DB 529 CSDPH 534

RESULT 10
 A48523
 margatoxin [validated] - scorpion (Centruroides margaritatus)
 N:Alternate names: potassium channel inhibitor Mgtx
 C:Species: Centruroides margaritatus
 C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 15-Sep-2000
 A:Accession: A48523
 R:Garcia-Calvo, M.; Leonard, R.J.; Novick, J.; Stevens, S.P.; Schmalhofer, W.; Kaczorowski
 J. Biol. Chem. 268, 18866-18874, 1993
 A>Title: Purification, characterization, and biosynthesis of margatoxin, a component of C
 A:Reference number: A48523; PMID:93366802; PMID:8360176
 A:Accession: A48523
 A:Molecule type: protein
 A:Residues: 1-39 <GAR>
 R:Johnson, B.A.; Stevens, S.P.; Williamson, J.M.
 submitted to the Brookhaven Protein Data Bank, December 1994
 A:Reference number: A66180; PDB:1MTX
 A:Contents: annotation; conformation and disulfide bond assignments by (1)H-NMR, residue
 R:Johnson, B.A.; Stevens, S.P.; Williamson, J.M.
 Biochemistry 33, 15061-15070, 1994
 A>Title: Determination of the three-dimensional structure of margatoxin by (1)H, (13)C,
 A:Reference number: A56028; PMID:95092763; PMID:7999764
 A:Contents: annotation; conformation and disulfide bond assignments by (1)H-, (13)C-, an
 C:Superfamily: kalitoxin
 C:Keywords: neurotoxin; potassium channel inhibitor; venom
 F:7-29,13-34,17-36/Disulfide bonds: #status experimental

Query Match 73.8%; Score 31; DB 1; Length 39;
 Best Local Similarity 66.7%; Pred. No. 71;
 Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CXKXP 6
 DB 34 CKCYP 39

RESULT 11
 JY0044
 hypothetical 7.5K protein (lepb 5' region) - Escherichia coli
 C:Species: Escherichia coli
 C:Date: 31-Mar-1990 #sequence_revision 31-Mar-1990 #text_change 18-Jun-1993
 A:Accession: JY0044
 R:Elkins, M.F.; Barhart, C.P.
 J. Bacteriol. 171, 5443-5451, 1989
 A>Title: Nucleotide sequence and regulation of the Escherichia coli gene for ferrientero
 A:Reference number: A91900; PMID:9008779; PMID:2529253
 A:Accession: JY0044
 A:Molecule type: DNA
 A:Residues: 1-68 <ELK>
 A:Experimental source: strain K12
 C:Genetics:
 A:Map position: 14 min

Query Match 73.8%; Score 31; DB 2; Length 68;
 Best Local Similarity 80.0%; Pred. No. 1.1e+02;
 Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 3 CXKXP 7
 DB 35 CSDPH 39

RESULT 12
 A35666
 transcription activator Krox-24 88K - mouse (fragment)
 C:Species: Mus musculus (house mouse)
 C:Date: 21-Sep-1990 #sequence_revision 21-Sep-1990 #text_change 02-Jul-1998
 A:Accession: A35666
 R:Lemaire, P.; Vesque, C.; Schmitt, J.; Stunnenberg, H.; Frank, R.; Charnay, P.
 Mol. Cell. Biol. 10, 3456-3467, 1990
 A>Title: The serum-inducible mouse gene Krox-24 encodes a sequence-specific transcription

A:Reference number: A35666; UID:90287135; PMID:2113174
 A:Accession: A35666
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-87 <LEM>
 A:Cross-references: GB:M38174

Query Match 73.8%; Score 31; DB 2; Length 87;
 Best Local Similarity 66.7%; Pred. No. 1.3e+02;
 Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CXKXPH 6
 DB 18 CPTCPH 23

RESULT 13

amicyanin precursor - Methylobacterium extorquens (strain AM1)
 C:Species: Methylobacterium extorquens
 C:Date: 04-Dec-1986 #sequence_revision 21-Jul-1995 #text_change 20-Apr-2000
 C:Accession: A56621; A00295
 R:Chisoberdov, A.Y.; Tsygankov, Y.D.; Lidetrom, M.E.
 DNA Seq. 2, 53-55, 1991
 A:Title: Nucleotide sequence of the amicyanin gene from Methylobacterium extorquens AM1.
 A:Reference number: A56621; UID:92199244; PMID:1802036
 A:Accession: A56621

A:Molecule type: DNA
 A:Residues: 1-119 <CHI>
 A:Cross-references: GB:M57963; NID:G150014; PIDN:AA68895.1; PID:G150016
 A:Note: sequence modified after extraction from NCBI backbone
 A:Note: the authors translated the codon CAC for residue 70 as Asn
 A:Note: sequence extracted from NCBI backbone (NCBIN:89409, NCBI:89412)

R:Amber, R.P.; Tobari, J.
 Biochem. J. 232, 451-457, 1985
 A:Title: The primary structures of Pseudomonas AM1 amicyanin and pseudoazurin. Two new B
 A:Reference number: A90327; UID:86130354; PMID:4091802
 A:Accession: A00295

A:Molecule type: protein
 A:Residues: 21-119 <AM>
 C:Comment: This species of Pseudomonas, isolated as an airborne contaminant, uses compou
 the true pseudomonads as well as methylobactrophs.

C:Superfamily: Plastocyanin
 C:Keywords: copper; electron transfer; metalloprotein; periplasmic space
 F:1-20/Domain: signal sequence #status predicted <Sig>
 F:21-119/Product: amicyanin #status experimental <Mat>
 F:67,106,109,112/Binding site: copper (His, Cys, His, Met) (type 1) #status predicted

Query Match 73.8%; Score 31; DB 1; Length 119;
 Best Local Similarity 80.0%; Pred. No. 1.7e+02;
 Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 3 CXKXPH 7
 DB 106 CTPHP 110

RESULT 14

amicyanin - Paracoccus denitrificans
 C:Species: Paracoccus denitrificans
 C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 20-Apr-2000
 C:Accession: S12972

R:van Spanning, R.J.M.; Wansell, C.W.; Reijnders, W.N.M.; Oltmann, L.F.; Stoutamer, A.H
 FBS Lett. 275, 217-220, 1990
 A:Title: Mutagenesis of the gene encoding amicyanin of Paracoccus denitrificans and the

A:Reference number: S12971; UID:91085564; PMID:2261991
 A:Accession: S12972
 A:Status: preliminary

A:Molecule type: DNA
 A:Residues: 1-131 <SPA>

A:Cross-references: EMBL:X55665; NID:G45458; PIDN:CAA39199.1; PID:G45460
 C:Superfamily: plastocyanin

C:Keywords: copper; electron transfer; metalloprotein
 F:79,118,121,124/Binding site: copper (His, Cys, His, Met) (type 1) #status experimental

Query Match 73.8%; Score 31; DB 1; Length 131;
 Best Local Similarity 80.0%; Pred. No. 1.8e+02;
 Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 3 CXKXPH 7
 DB 118 CTPHP 122

RESULT 15

amicyanin - Thiobacillus versutus
 A:Accession: A23706
 N:Alternate names: amine dehydrogenase amicyanin component
 C:Species: Thiobacillus versutus
 C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
 C:Accession: S19732; A23706

R:Ubbink, M.; van Kleef, M.A.G.; Kleinjan, D.J.; Holtink, C.W.G.; Hulstema, F.; Beintema,
 Eur. J. Biochem. 202, 1003-1012, 1991
 A:Title: Cloning, sequencing and expression studies of the genes encoding amicyanin and
 A:Reference number: S19730; UID:92111471; PMID:1765062
 A:Accession: S19732

A:Status: preliminary
 A:Molecule type: DNA

A:Residues: 1-132 <UB>
 A:Cross-references: GB:M58001; NID:G154632; PIDN:AA50571.1; PID:G154635
 R:Van Beeuwen, J.; Van Bun, S.; Canters, G.W.; Lommen, A.; Chochoia, C.
 J. Biol. Chem. 266, 4869-4877, 1991

A:Title: The structural homology of amicyanin from Thiobacillus versutus to plant plastoc
 A:Reference number: A23706; UID:91161570; PMID:2002033
 A:Accession: A23706

A:Status: preliminary
 A:Molecule type: protein
 A:Residues: 28-132 <VAN>
 C:Superfamily: Plastocyanin

C:Keywords: copper; electron transfer; metalloprotein
 F:80,119,122,125/Binding site: copper (His, Cys, His, Met) (type 1) #status predicted

Query Match 73.8%; Score 31; DB 1; Length 132;
 Best Local Similarity 80.0%; Pred. No. 1.8e+02;
 Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 3 CXKXPH 7
 DB 119 CTPHP 123

Search completed: April 8, 2004, 11:18:04
 Job time : 10.5278 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 8, 2004, 11:00:44 ; Search time 5.8333 Seconds

(without alignments)
62.484 Million cell updates/sec

Title: US-09-753-139c-1

Sequence: 42
1 CXCXPDP 7

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_42.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No	Score	Query Match	Length	ID	Description
1	32	76.2	830	1 SRC2_HUMAN	Q14162 homo sapien
2	32	76.2	833	1 SRC2_MOUSE	P59222 mus musculus
3	32	76.2	870	1 SRC2_HUMAN	Q96966 homo sapien
4	32	76.2	1172	1 LMB3_HUMAN	Q13751 homo sapien
5	32	76.2	1376	1 VGL2_CCM4	P22432 murine coro
6	32	76.2	1376	1 VGL2_CCM4	Q02385 murine coro
7	32	76.2	1428	1 ATRN_MOUSE	Q9w660 mus musculus
8	32	76.2	1587	1 LMB3_HUMAN	Q9y6n6 homo sapien
9	32	76.2	1786	1 LMB1_HUMAN	P07942 homo sapien
10	32	76.2	1786	1 LMB1_HUMAN	P02469 mus musculus
11	31	73.8	31	1 SCRT2_MOUSE	P59870 mesobutrus
12	31	73.8	39	1 SCRT1_MOUSE	P59847 centurionide
13	31	73.8	39	1 SCRT1_MOUSE	P40755 centurionide
14	31	73.8	119	1 AMCY_METEX	P04172 methylact
15	31	73.8	131	1 AMCY_PARDE	P22365 paracoccus
16	31	73.8	132	1 AMCY_PARVE	P22365 paracoccus
17	31	73.8	235	1 YHBF_ARCST	P39133 bacillus su
18	31	73.8	242	1 SCNC_THRTT	O66188 thlobacilli
19	31	73.8	313	1 GAG_AVISN	P03342 avian splee
20	31	73.8	317	1 VN35_ROTBN	Q03348 bovine roca
21	31	73.8	317	1 VN35_ROTBU	Q03348 bovine roca
22	31	73.8	317	1 VN35_ROTTH	Q03348 bovine roca
23	31	73.8	317	1 VN35_ROTTH	Q03348 bovine roca
24	31	73.8	317	1 VN35_ROTTH	Q03348 bovine roca
25	31	73.8	317	1 VN35_ROTTH	Q03348 bovine roca
26	31	73.8	317	1 VN35_ROTTH	Q03348 bovine roca
27	31	73.8	317	1 VN35_ROTTH	Q03348 bovine roca
28	31	73.8	317	1 VN35_ROTTH	Q03348 bovine roca
29	31	73.8	317	1 VN35_ROTTH	Q03348 bovine roca
30	31	73.8	317	1 VN35_ROTTH	Q03348 bovine roca
31	31	73.8	317	1 VN35_ROTTH	Q03348 bovine roca
32	31	73.8	317	1 VN35_ROTTH	Q03348 bovine roca
33	31	73.8	317	1 VN35_ROTTH	Q03348 bovine roca

ALIGNMENTS

RESULT 1	ID	SRRC_HUMAN	STANDARD	PRT	830 AA.
AC	Q14162	043701			
DT	28-FEB-2003	(Rel. 41, Created)			
DT	28-FEB-2003	(Rel. 41, Last sequence update)			
DT	10-OCT-2003	(Rel. 42, Last annotation update)			
DE	Endothelial cells scavenger receptor precursor (Acetyl LDL receptor)				
DE	(Scavenger receptor class F member 1).				
GN	SCARF1 OR SRC2 OR KIA0149.				
OS	Homo sapiens (Human).				
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;				
OC	Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.				
OX	NCBI_TaxID=9606;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RC	TISSUE=Umbilical vein endothelial cells;				
RX	MEDLINE=98058897; PubMed=9395444;				
RA	Adachi H., Tsujimoto M., Arai H., Inoue K.;				
RT	"Expression cloning of a novel scavenger receptor from human				
RL	endothelial cells.";				
RN	J. Biol. Chem. 272:31217-31220(1997).				
RN	[2]				
RP	SEQUENCE FROM N.A.				
RX	MEDLINE=22086180; PubMed=11978792;				
RA	Adachi H., Tsujimoto M.;				
RT	"Characterization of the human gene encoding the scavenger receptor				
RL	expressed by endothelial cell and its regulation by a novel				
RN	transcription factor, endothelial zinc finger protein-2.";				
RN	J. Biol. Chem. 277:24014-24021(2002).				
RN	[3]				
RP	SEQUENCE FROM N.A.				
RC	TISSUE=Bone marrow;				
RX	MEDLINE=96127530; PubMed=8590280;				
RA	Nagase T., Seki N., Tanaka A., Ishikawa K.-I., Nomura N.;				
RT	"Prediction of the coding sequences of unidentified human genes. IV.				
RL	The coding sequences of 40 new genes (K1A0121-K1A0160) deduced by				
RN	analysis of cDNA clones from human cell line KG-1.";				
RN	DNA Res. 2:167-174(1995).				
RN	[4]				
RP	SEQUENCE FROM N.A.				
RC	TISSUE=Testis;				
RX	MEDLINE=22388257; PubMed=12477932;				
RA	Klausner R.L., Feingold R.A., Grouse L.H., Derge J.G.,				
RT	Shenmen C.M., Schuler G.D.,				
RL	Altechi S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,				
RN	Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,				
RA	Diachenko L., Maruina K., Farmer A.A., Rubin G.M., Hong L.,				
RT	Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.B.,				
RL	Brownstein M.J., Udell T.B., Toehiyuki S., Carrincci P., Prange C.,				
RN	Raha S.S., Loquellano N.A., Peters G.J., Adamson R.D., Mullany S.J.,				
RA	Boeck S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,				
RT	Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Philyk S.W.,				
RL	Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,				
RN	Fabry J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,				
RA	Whitling M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,				

34	31	73.8	502	1	P2BC_HUMAN	P48454 homo sapien
35	31	73.8	506	1	ABSA_MOUSE	P50428 mus musculus
36	31	73.8	508	1	LCK_HUMAN	P06239 homo sapien
37	31	73.8	513	1	P2BC_MOUSE	P48455 mus musculus
38	31	73.8	515	1	P2BB_MOUSE	P48453 mus musculus
39	31	73.8	518	1	P2B_NEUCR	Q05681 neurospora
40	31	73.8	521	1	P2BA_BOVIN	P48452 bos taurus
41	31	73.8	521	1	P2BA_HUMAN	Q08209 homo sapien
42	31	73.8	521	1	P2BA_MOUSE	P20652 mus musculus
43	31	73.8	524	1	P2BB_HUMAN	P16239 homo sapien
44	31	73.8	525	1	P2BB_RAT	P20651 rattus norv
45	31	73.8	530	1	P2B_EWENT	P48457 emericeella

RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Buterfield Y.S.N., Krzywinski M.I., Skalska U., Smalins D.E.,
 RA Schermer A., Schin J.E., Jones S.J.W., Marra M.A.,
 RT "Generation and initial analysis of more than 15,000 full-length
 human and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 CC -1- FUNCTION: Mediates the binding and degradation of acetylated low
 density lipoprotein (Ac-LDL). Mediates heterophilic interactions,
 CC suggesting a function as adhesion protein (By similarity).
 CC -1- SUBUNIT: Heterophilic interaction with SREC2 via its extracellular
 CC domain. The heterophilic interaction is suppressed by the presence
 CC of ligand such as Ac-LDL (By similarity).
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein (Potential).
 CC -1- TISSUE SPECIFICITY: Endothelial cells.
 CC -1- SIMILARITY: Contains 6 EGF-like domains.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; D86864; BAA24070.1; -;
 DR EMBL; AB052946; BAC02692.1; -;
 DR EMBL; D63483; BAA09770.1; -;
 DR EMBL; BC039735; AAB39735.1; -;
 DR HSSP; F01180; 2BN2.
 DR GeneW; HGNC:16820; SCARF1.
 DR MIM; 607873; -;
 DR GO; GO:0016021; C:integral to membrane; IDA.
 DR GO; GO:0010169; F:low-density lipoprotein binding; IDA.
 DR GO; GO:0004888; F:transmembrane receptor activity; TAS.
 DR GO; GO:0045192; F:low-density lipoprotein catabolism; TAS.
 DR GO; GO:0006898; P:receptor mediated endocytosis; TAS.
 DR InterPro; IPR006209; EGF like.
 DR InterPro; IPR009030; Grow_fac_recep.
 DR InterPro; IPR006210; IEGF_
 DR InterPro; IPR002049; Laminin_EGF.
 DR PRINTS; PR00011; EGF_LAMININ.
 DR SMART; SM00182; EGF_5.
 DR PROSITE; PS00022; EGF_1; 6.
 DR PROSITE; PS01186; EGF_2; 6.
 DR PROSITE; PS00026; EGF_3; 3.
 KM Cell adhesion; Receptor; Repeat; Signal; Transmembrane;
 KM EGF-like domain; Glycoprotein.
 FT SIGNAL 1 19 POTENTIAL.
 FT CHAIN 20 830 ENDOTHELIAL CELLS SCAVENGER RECEPTOR.
 FT DOMAIN 20 421 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 422 442 POTENTIAL.
 FT DOMAIN 443 830 CYTOPLASMIC (POTENTIAL).
 FT DOMAIN 53 87 EGF-LIKE 1.
 FT DOMAIN 95 130 EGF-LIKE 2.
 FT DOMAIN 155 191 EGF-LIKE 3.
 FT DOMAIN 215 249 EGF-LIKE 4.
 FT DOMAIN 302 339 EGF-LIKE 5.
 FT DOMAIN 351 382 EGF-LIKE 6.
 FT DOMAIN 476 620 PRO/SER-RICH.
 FT DOMAIN 622 798 GLY-RICH.
 FT DOMAIN 431 438 POLY-LEU.
 FT DOMAIN 57 69 POTENTIAL.
 FT DISULFID 63 75 POTENTIAL.
 FT DISULFID 77 86 POTENTIAL.
 FT DISULFID 99 111 POTENTIAL.
 FT DISULFID 105 118 POTENTIAL.
 FT DISULFID 120 129 POTENTIAL.
 FT DISULFID 159 172 POTENTIAL.
 FT DISULFID 165 179 POTENTIAL.
 FT DISULFID 181 190 POTENTIAL.
 FT DISULFID 219 230 POTENTIAL.
 FT DISULFID 225 237 POTENTIAL.

FT DISULFID 239 248 POTENTIAL.
 FT DISULFID 306 319 POTENTIAL.
 FT DISULFID 313 326 POTENTIAL.
 FT DISULFID 329 338 POTENTIAL.
 FT DISULFID 355 363 POTENTIAL.
 FT DISULFID 358 370 POTENTIAL.
 FT DISULFID 372 381 POTENTIAL.
 FT CARBOHYD 289 289 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 382 382 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 393 393 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CONFLICT 662 662 R -> W (IN REF. 3).
 SQ SEQUENCE 830 AA; 87430 MW; F560D9E1AA64D779 CRC64;
 Query Match 76.2%; Score 32; DB 1; Length 830;
 Best Local Similarity 66.7%; Pred. No. 2,1e+02;
 Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 Oy 1 CXCXPH 6
 Db 133 CACGPH 138
 RESULT 2
 ID SRC2_MOUSE STANDARD; PRT; 833 AA.
 AC P59222;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Scavenger receptor class F member 2 precursor (Scavenger receptor
 DE expressed by endothelial cells 2 protein) (SRBC-II).
 GN SCARF2 OR SREC2.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A., AND CHARACTERIZATION.
 RC STRAIN=C57BL/6J;
 RX MEDLINE=22267235; PubMed=12154095;
 RA Ishii J., Adachi H., Aoki J., Koizumi H., Tomita S., Suzuki T.,
 RA Tsujimoto M., Inoue K., Arai H.
 RT "SRBC-II, a new member of the scavenger receptor type F family,
 RT trans-interacts with SREC-I through its extracellular domain";
 RL J. Biol. Chem. 277:39696-39702(2002).
 CC -1- FUNCTION: Probable adhesion protein, which mediates homophilic and
 CC heterophilic interactions. In contrast to SCARF1, it poorly
 CC mediates the binding and degradation of acetylated low density
 CC lipoprotein (Ac-LDL).
 CC -1- SUBUNIT: Homophilic and heterophilic interaction via its
 CC extracellular domain. Interacts with SCARF1. The heterophilic
 CC interaction with SCARF1, which is stronger than the homophilic
 CC interaction with itself, is suppressed by the presence of SCARF1
 CC ligand such as Ac-LDL.
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein (Potential).
 CC -1- SIMILARITY: Contains 7 EGF-like domains.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; AF522197; AAN45862.1; -;
 DR MGD; MGI:1858430; Scarf2.
 DR GO; GO:0005044; F:scavenger receptor activity; IDA.
 DR GO; GO:0007157; P:heterophilic cell adhesion; IDA.
 DR InterPro; IPR006209; EGF like.
 DR InterPro; IPR006210; IEGF_
 DR InterPro; IPR002049; Laminin_EGF.
 DR PRINTS; PR00011; EGF_LAMININ.


```

FT DOMAIN 463 830 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 76 110 BGF-LIKE 1.
FT DOMAIN 122 153 BGF-LIKE 2.
FT DOMAIN 154 182 BGF-LIKE 3.
FT DOMAIN 183 212 BGF-LIKE 4.
FT DOMAIN 213 241 BGF-LIKE 5.
FT DOMAIN 242 270 BGF-LIKE 6.
FT DOMAIN 372 403 BGF-LIKE 7.
FT DOMAIN 652 851 PRO-RICH.
FT DISULFID 80 92 POTENTIAL.
FT DISULFID 86 98 POTENTIAL.
FT DISULFID 100 109 POTENTIAL.
FT DISULFID 126 134 POTENTIAL.
FT DISULFID 128 141 POTENTIAL.
FT DISULFID 143 152 POTENTIAL.
FT DISULFID 156 163 POTENTIAL.
FT DISULFID 158 170 POTENTIAL.
FT DISULFID 172 181 POTENTIAL.
FT DISULFID 185 193 POTENTIAL.
FT DISULFID 187 200 POTENTIAL.
FT DISULFID 202 211 POTENTIAL.
FT DISULFID 215 222 POTENTIAL.
FT DISULFID 217 229 POTENTIAL.
FT DISULFID 231 240 POTENTIAL.
FT DISULFID 244 251 POTENTIAL.
FT DISULFID 246 258 POTENTIAL.
FT DISULFID 260 269 POTENTIAL.
FT DISULFID 376 384 POTENTIAL.
FT DISULFID 379 391 POTENTIAL.
FT DISULFID 393 402 POTENTIAL.
FT CARBOHYD 83 83 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 310 310 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 365 365 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 403 403 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT VARIANT 777 777 D -> E (1n dBSNP:759611).
FT VARIANT 778 778 /FTID=VAR_015148.
FT VARIANT 778 778 V -> L (1n dBSNP:759612).
FT VARIANT 819 819 A -> G (1n dBSNP:874100).
FT VARIANT 819 819 /FTID=VAR_015150.
FT VARIANT 837 837 A -> G (1n dBSNP:874101).
FT CONFLICT 474 478 /FTID=VAR_015151.
FT CONFLICT 626 641 MISSING (IN REF. 2).
FT SEQUENCE 870 AA; 92479 MW; DCS735A508693DIF CRC64;

Query Match 76.2%; Score 32; DB 1; Length 870;
Best Local Similarity 66.7%; Pred. No. 2.2e+02;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

```

QY 1 CXCPH 6
DB 126 CSHPH 131

```

RESULT 4

```

LMB3 HUMAN
ID LMB3 HUMAN STANDARD; PRT: 1172 AA.
AC Q13751; O14947; Q14733; Q9UUK4; Q9UUL1;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Laminin beta-3 chain precursor (Laminin 5 beta 3) (Laminin B1k chain)
DE (Kalinin B1 chain).
GN LAMB3.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxId=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=95293372; PubMed=7774918;

```

```

RA Pulkkinen L., Gerecke D.R., Christiano A.M., Wagman D.W.,
RA Burgeson R.E., Uitto J.;
RT "Cloning of the beta 3 chain gene (LAMB3) of human laminin 5, a
RT candidate gene in junctional epidermolysis bullosa.";
RL Genomics 25:192-198(1995).
RN
RP SEQUENCE FROM N.A., AND SEQUENCE OF 172-190.
RX MEDLINE=94209274; PubMed=7512558;
RA Gerecke D.R., Wagman D.W., Champiaud M.F., Burgeson R.E.;
RT "The complete primary structure for a novel laminin chain, the
RT laminin B1k chain.";
RL J. Biol. Chem. 269:11073-11080(1994).
RN
RP SEQUENCE FROM N.A.
RX MEDLINE=21221101; PubMed=11296269;
RA Robbins P.B., Lin Q., Goodnough J.B., Tian H., Chen X., Khavari P.A.;
RT "In vivo restoration of laminin 5 beta 3 expression and function in
RT junctional epidermolysis bullosa.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:5193-5198(2001).
RN
RP SEQUENCE FROM N.A.
RX TISSUE-Epidermis;
RA MEDLINE=96015057; PubMed=8530036;
RA Morishima Y., Ariyama T., Yamashita K., Abe T., Ueda E., Yasuno H.,
RA Inazawa J.;
RT "Chromosomal loci of 50 human keratinocyte CDNA's assigned by
RT fluorescence in situ hybridization.";
RL Genomics 28:273-279(1995).
RN
RP SEQUENCE FROM N.A.
RA Grafham D.;
RT Submitted (APR-1999) to the EMBL/Genbank/DBJ databases.
RN
RP VARIANT JEB LEU-679.
RX MEDLINE=96055522; PubMed=7550237;
RA Pulkkinen L., McGrath J.A., Christiano A.M., Uitto J.;
RT "Detection of sequence variants in the gene encoding the beta 3 chain
RT of laminin 5 (LAMB3).";
RL Hum. Mutat. 6:77-84(1995).
RN
RP VARIANT GABEB LYS-210.
RX MEDLINE=99068967; PubMed=9767254;
RA Mellero J.E., Bady R.A.J., Aeherton D.J., Lake B.D., McGrath J.A.;
RT "E210K mutation in the gene encoding the beta3 chain of laminin-5
RT (LAMB3) is predictive of a phenotype of generalized atrophic benign
RT epidermolysis bullosa.";
RL Br. J. Dermatol. 139:325-331(1998).
RN
RP -1- FUNCTION: Binding to cells via a high affinity receptor, laminin
RP is thought to mediate the attachment, migration and organization
RP of cells into tissues during embryonic development by interacting
RP with other extracellular matrix components.
RN
RP -1- SUBUNIT: Laminin is a complex glycoprotein, consisting of three
RP different polypeptide chains (alpha, beta, gamma), which are bound
RP to each other by disulfide bonds into a cross-shaped molecule
RP comprising one long and three short arms with globules at each
RP end. The beta-3 chain is a subunit of laminin-5
RP (epiligrin/kalinin/nicein).
RN
RP -1- SUBCELLULAR LOCATION: Extracellular.
RN
RP -1- TISSUE SPECIFICITY: Found in the basement membranes (major
RP component).
RN
RP -1- DOMAIN: The alpha-helical domains I and II are thought to interact
RP with other laminin chains to form a coiled coil structure.
RN
RP -1- DOMAIN: Domain VI is globular.
RN
RP -1- DISEASE: Defects in LAMB3 are a cause of junctional epidermolysis
RP bullosa gravis (JEB) [MIM:226700]; also known as junctional
RP epidermolysis bullosa Herlitz-Pearson type. JEB is a blistering
RP disorder in skin that is characterized by a separation of basal
RP cells from the basement membrane due to a decreased number of
RP hemidesmosomes. Laminin-5 is missing from the basement membrane of
RP patients with the gravis form of epidermolysis bullosa.
RN
RP -1- DISEASE: Defects in LAMB3 are a cause of generalized atrophic
RP benign epidermolysis bullosa (GABEB) [MIM:22650]. This nonlethal
RP form of junctional epidermolysis bullosa is characterized by 11fe-

```


RX MEDLINE=91353381; PubMed=1966429;
RA Parker S.E., Buchmeier M.J.;
RT "RNA sequence analysis of the E2 genes of wildtype and
RL neuroattenuated mutants of MHV-4 reveals a hypervariable domain.";
ADV. Exp. Med. Biol. 276:395-402(1990).
CC -1- FUNCTION: THE PEPLIMER PROTEIN MEDIATES THE BINDING OF VIRIONS
CC TO THE HOST CELL RECEPTOR AND IS INVOLVED IN MEMBRANE FUSION
CC AND IN SYNCTIUM FORMATION.
CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
CC -1- SIMILARITY: NEARLY IDENTICAL TO THE E2 GLYCOPROTEINS FROM MHV-JHM
CC AND MHV-A59 STRAINS, EXCEPT FOR AN N-TERMINAL INSERTION.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL; M32789; AAA46456.1; -;
DR EMBL; S5114; AAB19590.1; -;
DR PIR; A33748; VG1H02.
DR InterPro; IPR002552; Corona_S2.
DR Pfam; PF01601; Corona_S2; 1_S2.
KM Glycoprotein; Envelope protein; Transmembrane; Signal.
FT SIGNAL 1 14
FT CHAIN 15 1376 E2 GLYCOPROTEIN.
FT CHAIN 15 769 SPIKE PROTEIN S1.
FT CHAIN 770 1376 SPIKE PROTEIN S2.
FT DOMAIN 15 1320 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 1321 1338 POTENTIAL.
FT DOMAIN 1339 1376 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 429 599 CYS-RICH.
FT CARBOHYD 31 31 IMPORTANT FOR THE NEUROVIRULENCE.
FT CARBOHYD 60 60 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 134 134 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 192 192 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 357 357 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 435 435 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 582 582 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 677 677 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 709 709 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 717 717 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 740 740 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 789 789 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 806 806 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 945 945 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1232 1232 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1242 1242 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1261 1261 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1277 1277 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1298 1298 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 1376 AA; 151882 MW; 88C01B97B252094E CRC64;
Query Match Score 32; DB 1; Length 1376;
Best Local Similarity 57.1%; Pred. No. 3.3e+02;
Matches 4; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
QY 1 CXCXPH 7
DB 532 CTCKNP 538
RESULT 6
VGL2_CVMUC STANDARD; PRT; 1376 AA.
AC Q02385;
DT 01-FEB-1994 (Rel. 28, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE E2 glycoprotein precursor (Spike glycoprotein) (Peplomer protein)

DE [Contains: Spike protein S1 (90B); Spike protein S2 (90A)].
GN S.
OS Murine coronavirus (strain JHMV / variant CL-2) (MHV) (Murine
OS hepatitis virus).
OS Viruses; ssRNA positive-strand viruses, no DNA stage; Nidovirales;
OC Coronaviridae; Coronavirinae.
OC NCBI_TaxID=33735;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92268864; PubMed=1316938;
RA Tauchi F., Ikeda T., Shida H.;
RT "Molecular cloning and expression of a spike protein of neurovirulent
RL murine coronavirus JHMV variant CL-2.";
J. Gen. Virol. 73:1065-1072(1992).
CC -1- FUNCTION: THE PEPLIMER PROTEIN MEDIATES THE BINDING OF VIRIONS
CC TO THE HOST CELL RECEPTOR AND IS INVOLVED IN MEMBRANE FUSION
CC AND IN SYNCTIUM FORMATION.
CC -1- SUBCELLULAR LOCATION: Type I membrane protein.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL; D10235; BAA01085.1; -;
DR PIR; J01534; J01534.
DR InterPro; IPR002552; Corona_S2.
DR Pfam; PF01601; Corona_S2; 1_S2.
KM Glycoprotein; Envelope protein; Transmembrane; Signal.
FT SIGNAL 1 14
FT CHAIN 15 1376 E2 GLYCOPROTEIN.
FT CHAIN 15 769 SPIKE PROTEIN S1.
FT CHAIN 770 1376 SPIKE PROTEIN S2.
FT DOMAIN 15 1320 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 1321 1338 POTENTIAL.
FT DOMAIN 1339 1376 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 429 599 CYS-RICH.
FT CARBOHYD 31 31 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 60 60 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 134 134 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 192 192 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 357 357 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 435 435 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 677 677 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 709 709 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 717 717 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 789 789 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 806 806 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 945 945 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1232 1232 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1242 1242 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1261 1261 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1277 1277 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1298 1298 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 1376 AA; 152041 MW; 98C30DD979F9E75 CRC64;
Query Match Score 32; DB 1; Length 1376;
Best Local Similarity 57.1%; Pred. No. 3.3e+02;
Matches 4; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
QY 1 CXCXPH 7
DB 532 CTCKNP 538
RESULT 7
ATRN_MOUSE STANDARD; PRT; 1428 AA.
AC Q9WU77;
DT 16-OCT-2001 (Rel. 40, Created)

DT 16-OCT-2001 (Rel. 40, last sequence update)
DR 28-FEB-2003 (Rel. 41, last annotation update)
DE Attractin precursor (Mahogany protein).
GN ATRN OR MGA OR MG.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
NC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
NCBI_TaxId:10090;
RN [1]
RP MEDLINE:99184159; PubMed:10086355;
RX Nagle D.L., McGrail S.H., Vitale J., Wolf E.A., Dussault B.J. Jr.,
RA DiRocco L., Holmgren L., Montano J., Bork P., Huszar D.,
RA Patchild-huntress V., Ge P., Kelly J., Ebeling C., Baldini L.,
RA Gluchart J., Burn P., Carlson G.A., Moore K.J.;
RT "The mahogany protein is a receptor involved in suppression of
obesity.";
RL Nature 398:148-152 (1999).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE:99184160; PubMed:10086356;
RA Gunn T.M., Miller K.A., He L., Hyman R.W., Davis R.W., Azarani A.,
RA Schlossman S.F., Duke-Cohan J.S., Barsh G.S.;
RT "The mouse mahogany locus encodes a transmembrane form of human
attractin.";
RL Nature 398:152-156 (1999).
RN [3]
RP FUNCTION.
RX MEDLINE:20578746; PubMed:11137996;
RA He L., Gunn T.M., Bouley D.M., Lu X.Y., Watson S.J., Schlossman S.F.,
RA Duke-Cohan J.S., Barsh G.S.;
RT "A biochemical function for attractin in agouti-induced pigmentation
and obesity.";
RL Nat. Genet. 27:40-47 (2001).
CC -1- FUNCTION: May play a role in melanocortin signaling pathways that
regulate energy homeostasis and hair color. Is a low-affinity
receptor for agouti.
CC -1- FUNCTION: Involved in the initial immune cell clustering during
inflammatory response and may regulate chemotactic activity of
chemokines (By similarity).
CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
CC -1- SIMILARITY: Contains 1 CUB domain.
CC -1- SIMILARITY: Contains 1 C-type lectin family domain.
CC -1- SIMILARITY: Contains 1 EGF-like domain.
CC -1- SIMILARITY: Contains 1 laminin EGF-like domain.
CC -1- CAUTION: Ref.1 sequence differs from that shown due to a
frameshift in position 1388.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AF116897; AAD20947.1; ALT_FRAME.
DR EMBL: AF119821; AAD25372.1; -.
DR EMBL: AF120318; AAD22476.1; -.
DR EMBL: AF120317; AAD22476.1; JOINED.
DR HSSP: P02468; IKDO.
DR MGD: MGI:1341628; Attn.
DR GO: GO:0004872; F:receptor activity; IPI.
DR GO: GO:0040014; P:regulation of body size; IMP.
DR InterPro: IPR000859; CUB.
DR InterPro: IPR006209; EGF like.
DR InterPro: IPR006652; Kelch rep.
DR InterPro: IPR002049; Laminin EGF.
DR InterPro: IPR001304; Laminin_C.
DR InterPro: IPR003659; Plexin-like.
DR InterPro: IPR002165; Plexin_repeat.
DR Pfam: PF00431; CUB; 1.
DR Pfam: PF01344; Kelch; 5.

DR Pfam: PF00053; laminin EGF; 1.
DR Pfam: PF00059; lectin_C; 1.
DR Pfam: PF01437; PSI; 4.
DR PRINTS: PR00011; EGF_LAMININ.
DR SMART: SM00034; CLECT; 1.
DR SMART: SM00042; CUB; 1.
DR SMART: SM00180; EGF_Lam; 1.
DR SMART: SM00423; PSI; 5.
DR PROSITE: PS50041; C_TYPE_LECTIN_2; 1.
DR PROSITE: PS01180; CUB; 1.
DR PROSITE: PS00022; EGF_1; 3.
DR PROSITE: PS01186; EGF_2; 1.
DR PROSITE: PS50026; EGF_3; 2.
DR PROSITE: PS01248; LAMININ_TYPE_EGF; 1.
KW Receptor; Laminin EGF-like domain; Glycoprotein;
KW Transmembrane; Inflammatory response; Lectin; Signal;
KW Alternative splicing.
FT SIGNAL 1 17
FT CHAIN 18 1428
FT DOMAIN 18 1278
FT TRANSMEM 1279 1299
FT DOMAIN 1300 1428
FT DOMAIN 100 128
FT DOMAIN 131 247
FT DOMAIN 794 918
FT DOMAIN 1076 1105
FT DOMAIN 45 76
FT DOMAIN 60 63
FT DISULFID 100 110
FT DISULFID 104 117
FT DISULFID 119 128
FT DISULFID 1062 1070
FT DISULFID 1064 1076
FT DISULFID 1079 1088
FT DISULFID 1091 1105
FT CARBOHYD 212 212
FT CARBOHYD 236 236
FT CARBOHYD 241 241
FT CARBOHYD 252 252
FT CARBOHYD 263 263
FT CARBOHYD 299 299
FT CARBOHYD 324 324
FT CARBOHYD 361 361
FT CARBOHYD 382 382
FT CARBOHYD 415 415
FT CARBOHYD 427 427
FT CARBOHYD 574 574
FT CARBOHYD 622 622
FT CARBOHYD 730 730
FT CARBOHYD 862 862
FT CARBOHYD 913 913
FT CARBOHYD 922 922
FT CARBOHYD 985 985
FT CARBOHYD 1042 1042
FT CARBOHYD 1053 1053
FT CARBOHYD 1072 1072
FT CARBOHYD 1197 1197
FT CARBOHYD 1205 1205
FT CARBOHYD 1249 1249
FT CARBOHYD 1258 1258
FT CONFLICT 6 6
FT CONFLICT 9 9
FT CONFLICT 20 20
FT CONFLICT 164 169
FT CONFLICT 255 255
FT CONFLICT 505 505
FT CONFLICT 833 833
FT CONFLICT 1075 1075
FT CONFLICT 1140 1140
FT CONFLICT 1171 1171
FT CONFLICT 1192 1192
FT CONFLICT 1201 1201
FT CONFLICT 1428 1428
AA: 158087
MW: D1EADBFPS5B8911 CRC64;


```

Query Match      76.2%; Score 32; DB 1; Length 1428;
Best Local Similarity 66.7%; Pred. No. 3.4e+02;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY      1 CXCPH 6
Db      306 CSCPH 311

RESULT 8
LMG3_HUMAN      STANDARD; PRT; 1587 AA.
AC      Q9Y6N6;
DT      28-FEB-2003 (Rel. 41, Created)
DT      28-FEB-2003 (Rel. 41, Last sequence update)
DT      15-MAR-2004 (Rel. 43, Last annotation update)
DE      Laminin gamma-3 chain precursor (laminin 12 gamma 3).
GN      LAMC3.
OS      Homo sapiens (Human).
OC      Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX      NCBI_Taxid=9606;
RN      [1]
RE      SEQUENCE FROM N.A.
RX      TISSUE=Placenta;
RX      MEDLINE=99242614; PubMed=10225960;
RA      Koch M., Olson P.F., Albus A., Jin W., Hunter D.D., Brunken W.J.,
RA      Burgess R.E., Champiand M.F.;
RT      "Characterization and expression of the laminin gamma3 chain: a novel,
RT      non-basement membrane-associated, laminin chain.";
RL      J. Cell Biol. 145:605-618 (1999).
CC      -1- FUNCTION: Binding to cells via a high affinity receptor, laminin
CC      is thought to mediate the attachment, migration and organization
CC      of cells into tissues during embryonic development by interacting
CC      with other extracellular matrix components.
CC      -1- SUBUNIT: Laminin is a complex glycoprotein, consisting of three
CC      different polypeptide chains (alpha, beta, gamma), which are bound
CC      to each other by disulfide bonds into a cross-shaped molecule
CC      comprising one long and three short arms with globules at each
CC      end. The gamma-3 chain is a subunit of laminin-12.
CC      -1- SUBCELLULAR LOCATION: Extracellular.
CC      -1- TISSUE SPECIFICITY: Broadly expressed in: skin, heart, lung, and
CC      the reproductive tract.
CC      -1- DOMAIN: The alpha-helical domains I and II are thought to interact
CC      with other laminin chains to form a coiled coil structure.
CC      -1- DOMAIN: Domain IV is globular.
CC      -1- SIMILARITY: Contains 1 laminin N-terminal domain.
CC      -1- SIMILARITY: Contains 11 laminin EGF-like domains.
CC      -1- SIMILARITY: Contains 1 laminin IV domain.
CC      -1- SIMILARITY: Contains 1 laminin IV domain.
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC      -----
DR      EMBL; AF041835; AAD36991.1; -.
DR      HSSP; P02468; TITLE.
DR      Genew; HGNC:6494; LAMC3.
DR      MIM; 604349; -.
DR      GO; GO:0005578; C:extracellular matrix; TAS.
DR      GO; GO:0016020; C:membrane; TAS.
DR      GO; GO:0005198; F:structural molecule activity; TAS.
DR      InterPro; IPR006209; EGF like.
DR      InterPro; IPR008212; Lam_N2.
DR      InterPro; IPR00034; Laminin_B.
DR      InterPro; IPR002049; Laminin_EGF.
DR      InterPro; IPR008211; LAMNT.
DR      Pfam; PF00052; laminin B; 1.
DR      Pfam; PF00053; laminin_EGF; 9.

```

```

DR      Pfam; PF00055; laminin_Nterm; 1.
DR      PRINTS; PR00011; EGF/LAMININ.
DR      PRODOM; PD02082; Lam_N2; 1.
DR      SMART; SM00180; EGF_Lam; 9.
DR      SMART; SM00136; LAMNT; 1.
DR      PROSITE; PS00022; EGF_1; 7.
DR      PROSITE; PS01186; EGF_2; 2.
DR      PROSITE; PS01248; LAMININ_TYPE_EGF; 10.
DR      PROSITE; PS01248; LAMININ_TYPE_EGF; 10.
KW      Glycoprotein; Basement membrane; Extracellular matrix; Coiled coil;
KW      Laminin EGF-like domain; Cell adhesion; Repeat; Signal.
FT      SIGNAL      1..19
FT      CHAIN      20..1587
FT      DOMAIN      20..270
FT      DOMAIN      271..326
FT      DOMAIN      327..382
FT      DOMAIN      383..429
FT      DOMAIN      430..479
FT      DOMAIN      480..489
FT      DOMAIN      490..572
FT      DOMAIN      573..706
FT      DOMAIN      707..754
FT      DOMAIN      755..809
FT      DOMAIN      810..865
FT      DOMAIN      866..916
FT      DOMAIN      917..964
FT      DOMAIN      965..1013
FT      DOMAIN      1014..1587
FT      DOMAIN      1071..1141
FT      DOMAIN      1200..1229
FT      DOMAIN      1424..1504
FT      DOMAIN      1535..1579
FT      SITE      1059..1061
FT      CARBOHYD      87
FT      CARBOHYD      119..119
FT      CARBOHYD      295
FT      CARBOHYD      328..328
FT      CARBOHYD      631..631
FT      CARBOHYD      837
FT      CARBOHYD      980
FT      CARBOHYD      1185
FT      CARBOHYD      1518
SQ      SEQUENCE      1587 AA; 172051 MW; 3C6B609B5F203319 CRC64;

Query Match      76.2%; Score 32; DB 1; Length 1587;
Best Local Similarity 66.7%; Pred. No. 3.8e+02;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY      1 CXCPH 6
Db      887 CSCPH 892

RESULT 9
LMB1_HUMAN      STANDARD; PRT; 1786 AA.
AC      P07942;
DT      01-AUG-1988 (Rel. 08, Created)
DT      01-AUG-1988 (Rel. 08, Last sequence update)
DT      15-MAR-2004 (Rel. 43, Last annotation update)
DE      Laminin beta-1 chain precursor (laminin B1 chain).
GN      LAMB1.
OS      Homo sapiens (Human).
OC      Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX      NCBI_Taxid=9606;
RN      [1]
RE      SEQUENCE FROM N.A.
RX      MEDLINE=90368768; PubMed=1975589;
RX      Vuciteenah R., Chow L.T., Tryggvason K.;
RT      "Structure of the human laminin B1 chain gene.";
RL      J. Biol. Chem. 265:15611-15616 (1990).
RN      [2]
RE      SEQUENCE FROM N.A.

```



```

FT DOMAIN 1431 1786 DOMAIN I.
FT DOMAIN 1216 1315 COILED COIL (POTENTIAL).
FT DOMAIN 1353 1388 COILED COIL (POTENTIAL).
FT DOMAIN 1442 1781 COILED COIL (POTENTIAL).
FT DISULFID 271 280 BY SIMILARITY.
FT DISULFID 273 298 BY SIMILARITY.
FT DISULFID 300 309 BY SIMILARITY.
FT DISULFID 312 332 BY SIMILARITY.
FT DISULFID 335 344 BY SIMILARITY.
FT DISULFID 337 362 BY SIMILARITY.
FT DISULFID 365 374 BY SIMILARITY.
FT DISULFID 377 395 BY SIMILARITY.
FT DISULFID 398 411 BY SIMILARITY.
FT DISULFID 400 426 BY SIMILARITY.
FT DISULFID 428 437 BY SIMILARITY.
FT DISULFID 440 455 BY SIMILARITY.
FT DISULFID 458 472 BY SIMILARITY.
FT DISULFID 460 479 BY SIMILARITY.
FT DISULFID 481 490 BY SIMILARITY.
FT DISULFID 493 507 BY SIMILARITY.
FT DISULFID 773 785 BY SIMILARITY.
FT DISULFID 775 792 BY SIMILARITY.
FT DISULFID 794 803 BY SIMILARITY.
FT DISULFID 806 818 BY SIMILARITY.
FT DISULFID 821 833 BY SIMILARITY.
FT DISULFID 823 840 BY SIMILARITY.
FT DISULFID 842 851 BY SIMILARITY.
FT DISULFID 854 864 BY SIMILARITY.
FT DISULFID 867 876 BY SIMILARITY.
FT DISULFID 869 883 BY SIMILARITY.
FT DISULFID 886 895 BY SIMILARITY.
FT DISULFID 898 914 BY SIMILARITY.
FT DISULFID 917 933 BY SIMILARITY.
FT DISULFID 919 944 BY SIMILARITY.
FT DISULFID 946 955 BY SIMILARITY.

```

```

Query Match Score 32; DB 1; Length 1786;
Best Local Similarity 66.7%; Pred. No. 4.2e+02;
Matches 4; Conservative 0; Mismatch 2; Indels 0; Gaps 0;

```

```

OY 1 CXCPH 6
Db 529 CSCRPH 534

```

```

RESULT 10
LMB1_MOUSE STANDARD; PRT; 1786 AA.
AC P02469;
DT 21-JUL-1986 (Rel. 01, Created)
DT 01-JUL-1989 (Rel. 11, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Laminin beta-1 chain precursor (Laminin B1 chain).
GN LAMB1-1 OR LAMB-1.
OS Mus musculus (Mouse).
OC Euteleostomi; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Euteleostomi; Rodentia; Sciurognathi; Muridae; Mus.
OX NCBI_Taxid=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=87147212; PubMed=3493487;
RA Sasaki M., Kato S., Kohno K., Martin G.R., Yamada Y.;
RT "Sequence of the cDNA encoding the laminin B1 chain reveals a
RL multidomain protein containing cysteine-rich repeats.";
RN Proc. Natl. Acad. Sci. U.S.A. 84:935-939(1987).
RP [2]
RP SEQUENCE OF 1292-1786 FROM N.A.
RX MEDLINE=85051302; PubMed=6209134;
RA Barlow D.P., Green N.M., Kurkinen M., Hogan B.L.M.;
RT "Sequencing of laminin B chain cDNAs reveals C-terminal regions of
RL coiled-coil alpha-helix.";
RN EMBO J. 3:2355-2362(1984).
RN [3]

```

```

RP SEQUENCE OF 165-172, 539-547 AND 712-719.
RC STRAIN=BAH/c; TISSUE=Endothelial cells;
RX MEDLINE=97363207; PubMed=9219532;
RA Frieser M., Noeckel H., Pausch F., Roeder C., Hahn A., Deutzmann R.,
RA Sorokin L.M.;
RT "Cloning of the mouse laminin alpha 4 cDNA. Expression in a subset of
RT endothelium.";
RL Eur. J. Biochem. 246:727-735(1997).
CC -1- FUNCTION: Binding to cells via a high affinity receptor, laminin
CC is thought to mediate the attachment, migration and organization
CC of cells into tissues during embryonic development by interacting
CC with other extracellular matrix components.
CC -1- SUBUNIT: Laminin is a complex glycoprotein, consisting of three
CC different polypeptide chains (alpha, beta, gamma), which are bound
CC to each other by disulfide bonds into a cross-shaped molecule
CC comprising one long and three short arms with globules at each
CC end. The beta-1 chain is a subunit of laminin-1 (EHS laminin),
CC laminin-2 (merosin), and laminin-6 (K-laminin).
CC -1- SUBCELLULAR LOCATION: Extracellular.
CC -1- TISSUE SPECIFICITY: Found in the basement membranes (major
CC component).
CC -1- SIMILARITY: Contains 1 laminin N-terminal domain.
CC -1- SIMILARITY: Contains 13 laminin EGF-like domains.
CC -1- SIMILARITY: Contains 1 laminin IV domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by, and for commercial
CC entities requires a license agreement (See http://www.ebi.ac.uk/announcements
CC or send an email to license@ebi.ac.uk).
CC -----
DR EMBL; M15525; AAA39407.1; ALT_INIT.
DR EMBL; X05212; CAA28839.1; -.
DR PIR; A26413; MAMSB1.
DR HSSP; P02468; LKLO.
DR MGI; MGI:96743; Lamb1-1.
DR InterPro; IPR006209; EGF_Like.
DR InterPro; IPR008211; LamNT.
DR Pfam; PF00053; laminin_EGF_13.
DR Pfam; PF00055; laminin_Nterm; 1.
DR PRINTS; PR00011; EGFLAMININ.
DR SMART; SM00180; EGF_Lam; 11.
DR SMART; SM0136; LamNT; 1.
DR PROSITE; PS00022; EGF_1; 9.
DR PROSITE; PS01186; EGF_2; 2.
DR PROSITE; PS01248; LAMININ_TYPE_EGF; 11.
DR GlycoProtEx; G01248; LAMININ; Basement membrane; Extracellular matrix; Coiled coil;
KW Laminin EGF-like domain; Cell adhesion; Repeat; Signal.
FT SIGNAL 1 21
FT CHAIN 22 1786 LAMININ BETA-1 CHAIN.
FT DOMAIN 22 270 LAMININ N-TERMINAL (DOMAIN VII).
FT DOMAIN 271 334 LAMININ EGF-LIKE 1.
FT DOMAIN 335 397 LAMININ EGF-LIKE 2.
FT DOMAIN 398 457 LAMININ EGF-LIKE 3.
FT DOMAIN 458 509 LAMININ EGF-LIKE 4.
FT DOMAIN 510 540 LAMININ EGF-LIKE 5 (INCOMPLETE).
FT DOMAIN 541 572 LAMININ DOMAIN IV.
FT DOMAIN 573 620 LAMININ EGF-LIKE 6.
FT DOMAIN 621 666 LAMININ EGF-LIKE 7.
FT DOMAIN 667 716 LAMININ EGF-LIKE 8.
FT DOMAIN 717 775 LAMININ EGF-LIKE 9.
FT DOMAIN 776 827 LAMININ EGF-LIKE 10.
FT DOMAIN 828 883 LAMININ EGF-LIKE 11.
FT DOMAIN 884 933 LAMININ EGF-LIKE 12.
FT DOMAIN 934 1084 LAMININ EGF-LIKE 13.
FT DOMAIN 1085 1131 LAMININ EGF-LIKE 14.
FT DOMAIN 1132 1178 LAMININ EGF-LIKE 15.
FT DOMAIN 1179 1226 LAMININ EGF-LIKE 16.
FT DOMAIN 1227 1274 LAMININ EGF-LIKE 17.
FT DOMAIN 1275 1322 LAMININ EGF-LIKE 18.
FT DOMAIN 1323 1370 LAMININ EGF-LIKE 19.
FT DOMAIN 1371 1418 LAMININ EGF-LIKE 20.
FT DOMAIN 1419 1466 LAMININ EGF-LIKE 21.
FT DOMAIN 1467 1514 LAMININ EGF-LIKE 22.
FT DOMAIN 1515 1562 LAMININ EGF-LIKE 23.
FT DOMAIN 1563 1610 LAMININ EGF-LIKE 24.
FT DOMAIN 1611 1658 LAMININ EGF-LIKE 25.
FT DOMAIN 1659 1706 LAMININ EGF-LIKE 26.
FT DOMAIN 1707 1754 LAMININ EGF-LIKE 27.
FT DOMAIN 1755 1802 LAMININ EGF-LIKE 28.
FT DOMAIN 1803 1850 LAMININ EGF-LIKE 29.
FT DOMAIN 1851 1898 LAMININ EGF-LIKE 30.
FT DOMAIN 1899 1946 LAMININ EGF-LIKE 31.
FT DOMAIN 1947 1994 LAMININ EGF-LIKE 32.
FT DOMAIN 1995 2042 LAMININ EGF-LIKE 33.
FT DOMAIN 2043 2090 LAMININ EGF-LIKE 34.
FT DOMAIN 2091 2138 LAMININ EGF-LIKE 35.
FT DOMAIN 2139 2186 LAMININ EGF-LIKE 36.
FT DOMAIN 2187 2234 LAMININ EGF-LIKE 37.
FT DOMAIN 2235 2282 LAMININ EGF-LIKE 38.
FT DOMAIN 2283 2330 LAMININ EGF-LIKE 39.
FT DOMAIN 2331 2378 LAMININ EGF-LIKE 40.
FT DOMAIN 2379 2426 LAMININ EGF-LIKE 41.
FT DOMAIN 2427 2474 LAMININ EGF-LIKE 42.
FT DOMAIN 2475 2522 LAMININ EGF-LIKE 43.
FT DOMAIN 2523 2570 LAMININ EGF-LIKE 44.
FT DOMAIN 2571 2618 LAMININ EGF-LIKE 45.
FT DOMAIN 2619 2666 LAMININ EGF-LIKE 46.
FT DOMAIN 2667 2714 LAMININ EGF-LIKE 47.
FT DOMAIN 2715 2762 LAMININ EGF-LIKE 48.
FT DOMAIN 2763 2810 LAMININ EGF-LIKE 49.
FT DOMAIN 2811 2858 LAMININ EGF-LIKE 50.
FT DOMAIN 2859 2906 LAMININ EGF-LIKE 51.
FT DOMAIN 2907 2954 LAMININ EGF-LIKE 52.
FT DOMAIN 2955 3002 LAMININ EGF-LIKE 53.
FT DOMAIN 3003 3050 LAMININ EGF-LIKE 54.
FT DOMAIN 3051 3098 LAMININ EGF-LIKE 55.
FT DOMAIN 3099 3146 LAMININ EGF-LIKE 56.
FT DOMAIN 3147 3194 LAMININ EGF-LIKE 57.
FT DOMAIN 3195 3242 LAMININ EGF-LIKE 58.
FT DOMAIN 3243 3290 LAMININ EGF-LIKE 59.
FT DOMAIN 3291 3338 LAMININ EGF-LIKE 60.
FT DOMAIN 3339 3386 LAMININ EGF-LIKE 61.
FT DOMAIN 3387 3434 LAMININ EGF-LIKE 62.
FT DOMAIN 3435 3482 LAMININ EGF-LIKE 63.
FT DOMAIN 3483 3530 LAMININ EGF-LIKE 64.
FT DOMAIN 3531 3578 LAMININ EGF-LIKE 65.
FT DOMAIN 3579 3626 LAMININ EGF-LIKE 66.
FT DOMAIN 3627 3674 LAMININ EGF-LIKE 67.
FT DOMAIN 3675 3722 LAMININ EGF-LIKE 68.
FT DOMAIN 3723 3770 LAMININ EGF-LIKE 69.
FT DOMAIN 3771 3818 LAMININ EGF-LIKE 70.
FT DOMAIN 3819 3866 LAMININ EGF-LIKE 71.
FT DOMAIN 3867 3914 LAMININ EGF-LIKE 72.
FT DOMAIN 3915 3962 LAMININ EGF-LIKE 73.
FT DOMAIN 3963 4010 LAMININ EGF-LIKE 74.
FT DOMAIN 4011 4058 LAMININ EGF-LIKE 75.
FT DOMAIN 4059 4106 LAMININ EGF-LIKE 76.
FT DOMAIN 4107 4154 LAMININ EGF-LIKE 77.
FT DOMAIN 4155 4202 LAMININ EGF-LIKE 78.
FT DOMAIN 4203 4250 LAMININ EGF-LIKE 79.
FT DOMAIN 4251 4298 LAMININ EGF-LIKE 80.
FT DOMAIN 4299 4346 LAMININ EGF-LIKE 81.
FT DOMAIN 4347 4394 LAMININ EGF-LIKE 82.
FT DOMAIN 4395 4442 LAMININ EGF-LIKE 83.
FT DOMAIN 4443 4490 LAMININ EGF-LIKE 84.
FT DOMAIN 4491 4538 LAMININ EGF-LIKE 85.
FT DOMAIN 4539 4586 LAMININ EGF-LIKE 86.
FT DOMAIN 4587 4634 LAMININ EGF-LIKE 87.
FT DOMAIN 4635 4682 LAMININ EGF-LIKE 88.
FT DOMAIN 4683 4730 LAMININ EGF-LIKE 89.
FT DOMAIN 4731 4778 LAMININ EGF-LIKE 90.
FT DOMAIN 4779 4826 LAMININ EGF-LIKE 91.
FT DOMAIN 4827 4874 LAMININ EGF-LIKE 92.
FT DOMAIN 4875 4922 LAMININ EGF-LIKE 93.
FT DOMAIN 4923 4970 LAMININ EGF-LIKE 94.
FT DOMAIN 4971 5018 LAMININ EGF-LIKE 95.
FT DOMAIN 5019 5066 LAMININ EGF-LIKE 96.
FT DOMAIN 5067 5114 LAMININ EGF-LIKE 97.
FT DOMAIN 5115 5162 LAMININ EGF-LIKE 98.
FT DOMAIN 5163 5210 LAMININ EGF-LIKE 99.
FT DOMAIN 5211 5258 LAMININ EGF-LIKE 100.
FT DOMAIN 5259 5306 LAMININ EGF-LIKE 101.
FT DOMAIN 5307 5354 LAMININ EGF-LIKE 102.
FT DOMAIN 5355 5402 LAMININ EGF-LIKE 103.
FT DOMAIN 5403 5450 LAMININ EGF-LIKE 104.
FT DOMAIN 5451 5498 LAMININ EGF-LIKE 105.
FT DOMAIN 5499 5546 LAMININ EGF-LIKE 106.
FT DOMAIN 5547 5594 LAMININ EGF-LIKE 107.
FT DOMAIN 5595 5642 LAMININ EGF-LIKE 108.
FT DOMAIN 5643 5690 LAMININ EGF-LIKE 109.
FT DOMAIN 5691 5738 LAMININ EGF-LIKE 110.
FT DOMAIN 5739 5786 LAMININ EGF-LIKE 111.
FT DOMAIN 5787 5834 LAMININ EGF-LIKE 112.
FT DOMAIN 5835 5882 LAMININ EGF-LIKE 113.
FT DOMAIN 5883 5930 LAMININ EGF-LIKE 114.
FT DOMAIN 5931 5978 LAMININ EGF-LIKE 115.
FT DOMAIN 5979 6026 LAMININ EGF-LIKE 116.
FT DOMAIN 6027 6074 LAMININ EGF-LIKE 117.
FT DOMAIN 6075 6122 LAMININ EGF-LIKE 118.
FT DOMAIN 6123 6170 LAMININ EGF-LIKE 119.
FT DOMAIN 6171 6218 LAMININ EGF-LIKE 120.
FT DOMAIN 6219 6266 LAMININ EGF-LIKE 121.
FT DOMAIN 6267 6314 LAMININ EGF-LIKE 122.
FT DOMAIN 6315 6362 LAMININ EGF-LIKE 123.
FT DOMAIN 6363 6410 LAMININ EGF-LIKE 124.
FT DOMAIN 6411 6458 LAMININ EGF-LIKE 125.
FT DOMAIN 6459 6506 LAMININ EGF-LIKE 126.
FT DOMAIN 6507 6554 LAMININ EGF-LIKE 127.
FT DOMAIN 6555 6602 LAMININ EGF-LIKE 128.
FT DOMAIN 6603 6650 LAMININ EGF-LIKE 129.
FT DOMAIN 6651 6698 LAMININ EGF-LIKE 130.
FT DOMAIN 6699 6746 LAMININ EGF-LIKE 131.
FT DOMAIN 6747 6794 LAMININ EGF-LIKE 132.
FT DOMAIN 6795 6842 LAMININ EGF-LIKE 133.
FT DOMAIN 6843 6890 LAMININ EGF-LIKE 134.
FT DOMAIN 6891 6938 LAMININ EGF-LIKE 135.
FT DOMAIN 6939 6986 LAMININ EGF-LIKE 136.
FT DOMAIN 6987 7034 LAMININ EGF-LIKE 137.
FT DOMAIN 7035 7082 LAMININ EGF-LIKE 138.
FT DOMAIN 7083 7130 LAMININ EGF-LIKE 139.
FT DOMAIN 7131 7178 LAMININ EGF-LIKE 140.
FT DOMAIN 7179 7226 LAMININ EGF-LIKE 141.
FT DOMAIN 7227 7274 LAMININ EGF-LIKE 142.
FT DOMAIN 7275 7322 LAMININ EGF-LIKE 143.
FT DOMAIN 7323 7370 LAMININ EGF-LIKE 144.
FT DOMAIN 7371 7418 LAMININ EGF-LIKE 145.
FT DOMAIN 7419 7466 LAMININ EGF-LIKE 146.
FT DOMAIN 7467 7514 LAMININ EGF-LIKE 147.
FT DOMAIN 7515 7562 LAMININ EGF-LIKE 148.
FT DOMAIN 7563 7610 LAMININ EGF-LIKE 149.
FT DOMAIN 7611 7658 LAMININ EGF-LIKE 150.
FT DOMAIN 7659 7706 LAMININ EGF-LIKE 151.
FT DOMAIN 7707 7754 LAMININ EGF-LIKE 152.
FT DOMAIN 7755 7802 LAMININ EGF-LIKE 153.
FT DOMAIN 7803 7850 LAMININ EGF-LIKE 154.
FT DOMAIN 7851 7898 LAMININ EGF-LIKE 155.
FT DOMAIN 7899 7946 LAMININ EGF-LIKE 156.
FT DOMAIN 7947 7994 LAMININ EGF-LIKE 157.
FT DOMAIN 7995 8042 LAMININ EGF-LIKE 158.
FT DOMAIN 8043 8090 LAMININ EGF-LIKE 159.
FT DOMAIN 8091 8138 LAMININ EGF-LIKE 160.
FT DOMAIN 8139 8186 LAMININ EGF-LIKE 161.
FT DOMAIN 8187 8234 LAMININ EGF-LIKE 162.
FT DOMAIN 8235 8282 LAMININ EGF-LIKE 163.
FT DOMAIN 8283 8330 LAMININ EGF-LIKE 164.
FT DOMAIN 8331 8378 LAMININ EGF-LIKE 165.
FT DOMAIN 8379 8426 LAMININ EGF-LIKE 166.
FT DOMAIN 8427 8474 LAMININ EGF-LIKE 167.
FT DOMAIN 8475 8522 LAMININ EGF-LIKE 168.
FT DOMAIN 8523 8570 LAMININ EGF-LIKE 169.
FT DOMAIN 8571 8618 LAMININ EGF-LIKE 170.
FT DOMAIN 8619 8666 LAMININ EGF-LIKE 171.
FT DOMAIN 8667 8714 LAMININ EGF-LIKE 172.
FT DOMAIN 8715 8762 LAMININ EGF-LIKE 173.
FT DOMAIN 8763 8810 LAMININ EGF-LIKE 174.
FT DOMAIN 8811 8858 LAMININ EGF-LIKE 175.
FT DOMAIN 8859 8906 LAMININ EGF-LIKE 176.
FT DOMAIN 8907 8954 LAMININ EGF-LIKE 177.
FT DOMAIN 8955 9002 LAMININ EGF-LIKE 178.
FT DOMAIN 9003 9050 LAMININ EGF-LIKE 179.
FT DOMAIN 9051 9098 LAMININ EGF-LIKE 180.
FT DOMAIN 9099 9146 LAMININ EGF-LIKE 181.
FT DOMAIN 9147 9194 LAMININ EGF-LIKE 182.
FT DOMAIN 9195 9242 LAMININ EGF-LIKE 183.
FT DOMAIN 9243 9290 LAMININ EGF-LIKE 184.
FT DOMAIN 9291 9338 LAMININ EGF-LIKE 185.
FT DOMAIN 9339 9386 LAMININ EGF-LIKE 186.
FT DOMAIN 9387 9434 LAMININ EGF-LIKE 187.
FT DOMAIN 9435 9482 LAMININ EGF-LIKE 188.
FT DOMAIN 9483 9530 LAMININ EGF-LIKE 189.
FT DOMAIN 9531 9578 LAMININ EGF-LIKE 190.
FT DOMAIN 9579 9626 LAMININ EGF-LIKE 191.
FT DOMAIN 9627 9674 LAMININ EGF-LIKE 192.
FT DOMAIN 9675 9722 LAMININ EGF-LIKE 193.
FT DOMAIN 9723 9770 LAMININ EGF-LIKE 194.
FT DOMAIN 9771 9818 LAMININ EGF-LIKE 195.
FT DOMAIN 9819 9866 LAMININ EGF-LIKE 196.
FT DOMAIN 9867 9914 LAMININ EGF-LIKE 197.
FT DOMAIN 9915 9962 LAMININ EGF-LIKE 198.
FT DOMAIN 9963 10010 LAMININ EGF-LIKE 199.
FT DOMAIN 10011 10058 LAMININ EGF-LIKE 200.
FT DOMAIN 10059 10106 LAMININ EGF-LIKE 201.
FT DOMAIN 10107 10154 LAMININ EGF-LIKE 202.
FT DOMAIN 10155 10202 LAMININ EGF-LIKE 203.
FT DOMAIN 10203 10250 LAMININ EGF-LIKE 204.
FT DOMAIN 10251 10298 LAMININ EGF-LIKE 205.
FT DOMAIN 10299 10346 LAMININ EGF-LIKE 206.
FT DOMAIN 10347 10394 LAMININ EGF-LIKE 207.
FT DOMAIN 10395 10442 LAMININ EGF-LIKE 208.
FT DOMAIN 10443 10490 LAMININ EGF-LIKE 209.
FT DOMAIN 10491 10538 LAMININ EGF-LIKE 210.
FT DOMAIN 10539 10586 LAMININ EGF-LIKE 211.
FT DOMAIN 10587 10634 LAMININ EGF-LIKE 212.
FT DOMAIN 10635 10682 LAMININ EGF-LIKE 213.
FT DOMAIN 10683 10730 LAMININ EGF-LIKE 214.
FT DOMAIN 10731 10778 LAMININ EGF-LIKE 215.
FT DOMAIN 10779 10826 LAMININ EGF-LIKE 216.
FT DOMAIN 10827 10874 LAMININ EGF-LIKE 217.
FT DOMAIN 10875 10922 LAMININ EGF-LIKE 218.
FT DOMAIN 10923 10970 LAMININ EGF-LIKE 219.
FT DOMAIN 10971 11018 LAMININ EGF-LIKE 220.
FT DOMAIN 11019 11066 LAMININ EGF-LIKE 221.
FT DOMAIN 11067 11114 LAMININ EGF-LIKE 222.
FT DOMAIN 11115 11162 LAMININ EGF-LIKE 223.
FT DOMAIN 11163 11210 LAMININ EGF-LIKE 224.
FT DOMAIN 11211 11258 LAMININ EGF-LIKE 225.
FT DOMAIN 11259 11306 LAMININ EGF-LIKE 226.
FT DOMAIN 11307 11354 LAMININ EGF-LIKE 227.
FT DOMAIN 11355 11402 LAMININ EGF-LIKE 228.
FT DOMAIN 11403 11450 LAMININ EGF-LIKE 229.
FT DOMAIN 11451 11498 LAMININ EGF-LIKE 230.
FT DOMAIN 11499 11546 LAMININ EGF-LIKE 231.
FT DOMAIN 11547 11594 LAMININ EGF-LIKE 232.
FT DOMAIN 11595 11642 LAMININ EGF-LIKE 233.
FT DOMAIN 11643 11690 LAMININ EGF-LIKE 234.
FT DOMAIN 11691 11738 LAMININ EGF-LIKE 235.
FT DOMAIN 11739 11786 LAMININ EGF-LIKE 236.
FT DOMAIN 11787 11834 LAMININ EGF-LIKE 237.
FT DOMAIN 11835 11882 LAMININ EGF-LIKE 238.
FT DOMAIN 11883 11930 LAMININ EGF-LIKE 239.
FT DOMAIN 11931 11978 LAMININ EGF-LIKE 240.
FT DOMAIN 11979 12026 LAMININ EGF-LIKE 241.
FT DOMAIN 12027 12074 LAMININ EGF-LIKE 242.
FT DOMAIN 12075 12122 LAMININ EGF-LIKE 243.
FT DOMAIN 12123 12170 LAMININ EGF-LIKE 244.
FT DOMAIN 12171 12218 LAMININ EGF-LIKE 245.
FT DOMAIN 12219 12266 LAMININ EGF-LIKE 246.
FT DOMAIN 12267 12314 LAMININ EGF-LIKE 247.
FT DOMAIN 12315 12362 LAMININ EGF-LIKE 248.
FT DOMAIN 12363 12410 LAMININ EGF-LIKE 249.
FT DOMAIN 12411 12458 LAMININ EGF-LIKE 250.
FT DOMAIN 12459 12506 LAMININ EGF-LIKE 251.
FT DOMAIN 12507 12554 LAMININ EGF-LIKE 252.
FT DOMAIN 12555 12602 LAMININ EGF-LIKE 253.
FT DOMAIN 12603 12650 LAMININ EGF-LIKE 254.
FT DOMAIN 12651 12698 LAMININ EGF-LIKE 255.
FT DOMAIN 12699 12746 LAMININ EGF-LIKE 256.
FT DOMAIN 12747 12794 LAMININ EGF-LIKE 257.
FT DOMAIN 12795 12842 LAMININ EGF-LIKE 258.
FT DOMAIN 12843 12890 LAMININ EGF-LIKE 259.
FT DOMAIN 12891 12938 LAMININ EGF-LIKE 260.
FT DOMAIN 12939 12986 LAMININ EGF-LIKE 261.
FT DOMAIN 12987 13034 LAMININ EGF-LIKE 262.
FT DOMAIN 13035 13082 LAMININ EGF-LIKE 263.
FT DOMAIN 13083 13130 LAMININ EGF-LIKE 264.
FT DOMAIN 13131 13178 LAMININ EGF-LIKE 265.
FT DOMAIN 13179 13226 LAMININ EGF-LIKE 266.
FT DOMAIN 13227 13274 LAMININ EGF-LIKE 267.
FT DOMAIN 13275 13322 LAMININ EGF-LIKE 268.
FT DOMAIN 13323 13370 LAMININ EGF-LIKE 269.
FT DOMAIN 13371 13418 LAMININ EGF-LIKE 270.
FT DOMAIN 13419 13466 LAMININ EGF-LIKE 271.
FT DOMAIN 13467 13514 LAMININ EGF-LIKE 272.
FT DOMAIN 13515 13562 LAMININ EGF-LIKE 273.
FT DOMAIN 13563 13610 LAMININ EGF-LIKE 274.
FT DOMAIN 13611 13658 LAMININ EGF-LIKE 275.
FT DOMAIN 13659 13706 LAMININ EGF-LIKE 276.
FT DOMAIN 13707 13754 LAMININ EGF-LIKE 277.
FT DOMAIN 13755 13802 LAMININ EGF-LIKE 278.
FT DOMAIN 13803 13850 LAMININ EGF-LIKE 279.
FT DOMAIN 13851 13898 LAMININ EGF-LIKE 280.
FT DOMAIN 13899 13946 LAMININ EGF-LIKE 281.
FT DOMAIN 13947 13994 LAMININ EGF-LIKE 282.
FT DOMAIN 13995 14042 LAMININ EGF-LIKE 283.
FT DOMAIN 14043 14090 LAMININ EGF-LIKE 284.
FT DOMAIN 14091 14138 LAMININ EGF-LIKE 285.
FT DOMAIN 14139 14186 LAMININ EGF-LIKE 286.
FT DOMAIN 14187 14234 LAMININ EGF-LIKE 287.
FT DOMAIN 14235 14282 LAMININ EGF-LIKE 288.
FT DOMAIN 14283 14330 LAMININ EGF-LIKE 289.
FT DOMAIN 14331 14378 LAMININ EGF-LIKE 290.
FT DOMAIN 14379 14426 LAMININ EGF-LIKE 291.
FT DOMAIN 14427 14474 LAMININ EGF-LIKE 292.
FT DOMAIN 14475 14522 LAMININ EGF-LIKE 293.
FT DOMAIN 14523 14570 LAMININ EGF-LIKE 294.
FT DOMAIN 14571 14618 LAMININ EGF-LIKE 295.
FT DOMAIN 14619 14666 LAMININ EGF-LIKE 296.
FT DOMAIN 14667 14714 LAMININ EGF-LIKE 297.
FT DOMAIN 14715 14762 LAMININ EGF-LIKE 298.
FT DOMAIN 14763 14810 LAMININ EGF-LIKE 299.
FT DOMAIN 14811 14858 LAMININ EGF-LIKE 300.
FT DOMAIN 14859 14906 LAMININ EGF-LIKE 301.
FT DOMAIN 14907 14954 LAMININ EGF-LIKE 302.
FT DOMAIN 14955 15002 LAMININ EGF-LIKE 303.
FT DOMAIN 15003 15050 LAMININ EGF-LIKE 304.
FT DOMAIN 15051 15098 LAMININ EGF-LIKE 305.
FT DOMAIN 15099 15146 LAMININ EGF-LIKE 306.
FT DOMAIN 15147 15194 LAMININ EGF-LIKE 307.
FT DOMAIN 15195 15242 LAMININ EGF-LIKE 308.
FT DOMAIN 15243 15290 LAMININ EGF-LIKE 309.
FT DOMAIN 15291 15338 LAMININ EGF-LIKE 310.
FT DOMAIN 15339 15386 LAMININ EGF-LIKE 311.
FT DOMAIN 15387 15434 LAMININ EGF-LIKE 312.
FT DOMAIN 15435 15482 LAMININ EGF-LIKE 313.
FT DOMAIN 15483 15530 LAMININ EGF-LIKE 314.
FT DOMAIN 15531 15578 LAMININ EGF-LIKE 315.
FT DOMAIN 15579 15626 LAMININ EGF-LIKE 316.
FT DOMAIN 15627 15674 LAMININ EGF-LIKE 317.
FT DOMAIN 15675 15722 LAMININ EGF-LIKE 318.
FT DOMAIN 15723 15770 LAMININ EGF-LIKE 319.
FT DOMAIN 15771 15818 LAMININ EGF-LIKE 320.
FT DOMAIN 15819 15866 LAMININ EGF-LIKE 321.
FT DOMAIN 15867 15914 LAMININ EGF-LIKE 322.
FT DOMAIN 15915 15962 LAMININ EGF-LIKE 323.
FT DOMAIN 15963 16010 LAMININ EGF-LIKE 324.
FT DOMAIN 16011 16058 LAMININ EGF-LIKE 325.
FT DOMAIN 16059 16106 LAMININ EGF-LIKE 326.
FT DOMAIN 16107 16154 LAMININ EGF-LIKE 327.
FT DOMAIN 16155 16202 LAMININ EGF-LIKE 328.
FT DOMAIN 16203 16250 LAMININ EGF-LIKE 329.
FT DOMAIN 16251 16298 LAMININ EGF-LIKE 330.
FT DOMAIN 16299 16346 LAMININ EGF-LIKE 331.
FT DOMAIN 16347 16394 LAMININ EGF-LIKE 332.
FT DOMAIN 16395 16442 LAMININ EGF-LIKE 333.
FT DOMAIN 16443 16490 LAMININ EGF-LIKE 334.
FT DOMAIN 16491 16538 LAMININ EGF-LIKE 335.
FT DOMAIN 16539 16586 LAMININ EGF-LIKE 336.
FT DOMAIN 16587 16634 LAMININ EGF-LIKE 337.
FT DOMAIN 16635 16682 LAMININ EGF-LIKE 338.
FT DOMAIN 16683 16730 LAMININ EGF-LIKE 339.
FT DOMAIN 16731 16778 LAMININ EGF-LIKE 340.
FT DOMAIN 16779 16826 LAMININ EGF-LIKE 341.
FT DOMAIN 16827 16874 LAMININ EGF-LIKE 342.
FT DOMAIN 16875 16922 LAMININ EGF-LIKE 343.
FT DOMAIN 16923 16970 LAMININ EGF-LIKE 344.
FT DOMAIN 16971 17018 LAMININ EGF-LIKE 345.
FT DOMAIN 17019 17066 LAMININ EGF-LIKE 346.
FT DOMAIN 17067 17114 LAMININ EGF-LIKE 347.
FT DOMAIN 17115 17162 LAMININ EGF-LIKE 348.
FT DOMAIN 17163 17210 LAMININ EGF-LIKE 349.
FT DOMAIN 17211 17258 LAMININ EGF-LIKE 350.
FT DOMAIN 17259 17306 LAMININ EGF-LIKE 351.
FT DOMAIN 17307 17354 LAMININ EGF-LIKE 352.
FT DOMAIN 17355 17402 LAMININ EGF-LIKE 353.
FT DOMAIN 17403 17450 LAMININ EGF-LIKE 354.
FT DOMAIN 17451 17498 LAMININ EGF-LIKE 355.
FT DOMAIN 17499 17546 LAMININ EGF-LIKE 356.
FT DOMAIN 17547 17594 LAMININ EGF-LIKE 357.
FT DOMAIN 17595 17642 LAMININ EGF-LIKE 358.
FT DOMAIN 17643 17690 LAMININ EGF-LIKE 359.
FT DOMAIN 17691 17738 LAMININ EGF-LIKE 360.
FT DOMAIN 17739 17786 LAMININ EGF-LIKE 361.
FT DOMAIN 17787 17834 LAMININ EGF-LIKE 362.
FT DOMAIN 17835 17882 LAMININ EGF-LIKE 363.
FT DOMAIN 17883 17930 LAMININ EGF-LIKE 364.
FT DOMAIN 17931 17978 LAMININ EGF-LIKE 365.
FT DOMAIN 17979 18026 LAMININ EGF-LIKE 366.
FT DOMAIN 18027 18074 LAMININ EGF-LIKE 367.
FT DOMAIN 18075 18122 LAMININ EGF-LIKE 368.
FT DOMAIN 18123 18170 LAMININ EGF-LIKE 369.
FT DOMAIN 18171 18218 LAMININ EGF-LIKE 370.
FT DOMAIN 18219 
```

PT DOMAIN 1448 1778 COILED COIL (POTENTIAL).
 FT DISULFID 271 280 BY SIMILARITY.
 FT DISULFID 273 298 BY SIMILARITY.
 FT DISULFID 300 309 BY SIMILARITY.
 FT DISULFID 312 332 BY SIMILARITY.
 FT DISULFID 335 344 BY SIMILARITY.
 FT DISULFID 337 362 BY SIMILARITY.
 FT DISULFID 365 374 BY SIMILARITY.
 FT DISULFID 377 395 BY SIMILARITY.
 FT DISULFID 398 411 BY SIMILARITY.
 FT DISULFID 400 426 BY SIMILARITY.
 FT DISULFID 428 437 BY SIMILARITY.
 FT DISULFID 440 455 BY SIMILARITY.
 FT DISULFID 458 472 BY SIMILARITY.
 FT DISULFID 460 479 BY SIMILARITY.
 FT DISULFID 481 490 BY SIMILARITY.
 FT DISULFID 493 507 BY SIMILARITY.
 FT DISULFID 773 785 BY SIMILARITY.
 FT DISULFID 775 792 BY SIMILARITY.
 FT DISULFID 794 803 BY SIMILARITY.
 FT DISULFID 806 818 BY SIMILARITY.
 FT DISULFID 821 833 BY SIMILARITY.
 FT DISULFID 823 840 BY SIMILARITY.
 FT DISULFID 842 851 BY SIMILARITY.
 FT DISULFID 854 864 BY SIMILARITY.
 FT DISULFID 867 876 BY SIMILARITY.
 FT DISULFID 869 883 BY SIMILARITY.
 FT DISULFID 886 895 BY SIMILARITY.
 FT DISULFID 898 914 BY SIMILARITY.
 FT DISULFID 917 934 BY SIMILARITY.
 FT DISULFID 919 944 BY SIMILARITY.
 FT DISULFID 946 955 BY SIMILARITY.
 FT DISULFID 958 973 BY SIMILARITY.
 FT DISULFID 976 990 BY SIMILARITY.
 FT DISULFID 978 997 BY SIMILARITY.
 FT DISULFID 1000 1009 BY SIMILARITY.
 FT DISULFID 1012 1025 BY SIMILARITY.
 FT DISULFID 1084 1096 BY SIMILARITY.
 FT DISULFID 1086 1103 BY SIMILARITY.
 FT DISULFID 1105 1114 BY SIMILARITY.
 FT DISULFID 1117 1129 BY SIMILARITY.
 FT DISULFID 1132 1144 BY SIMILARITY.
 FT DISULFID 1134 1151 BY SIMILARITY.
 FT DISULFID 1153 1162 BY SIMILARITY.
 FT DISULFID 1165 1176 BY SIMILARITY.
 FT DISULFID 1179 1182 BY SIMILARITY.
 FT DISULFID 1182 1182 INTERCHAIN (PROBABLE).
 FT DISULFID 1785 1785 INTERCHAIN (PROBABLE).
 FT CARBOHYD 120 120 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 356 356 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 519 519 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 677 677 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 1041 1041 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 1195 1195 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 1279 1279 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 1336 1336 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 1343 1343 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 1487 1487 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 1533 1533 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 1542 1542 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 1643 1643 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 1643 1643 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CONFLICT 1531 1534 SGNA -> MEMB (IN REF. 2).
 FT CONFLICT 1749 1749 D -> N (IN REF. 2).
 SQ SEQUENCE 1786 AA; 196904 MW; 846671B7BF41A474 CRC64;

Query Match Best Local Similarity 76.2%; Score 32; DB 1; Length 1786;
 Pred. No. 4.2e+02; Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CXCXPH 6
 DB 529 CXCXPH 534

RESULT 11
 SCT2_MESTA ID SCT2_MESTA STANDARD; PRT; 31 AA.
 AC P59870;
 DT 10-OCT-2003 (Rel. 42, Created)
 DT 10-OCT-2003 (Rel. 42, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Tamapin-2.
 OS Mesobuthus tamulus (Eastern Indian scorpion) (Butus tamulus).
 OC Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Scorpiones;
 OC Butioidea; Butidae; Mesobuthus.
 OX NCBI_Taxid=34647;
 RN [1]
 RP SEQUENCE, MASS SPECTROMETRY, AND FUNCTION.
 RC TISSUE=Venom;
 RX MEDLINE=22336377; PubMed=12239213;
 RA Pedarzani P., D'hoedt D., Doorty K.B., Madsenworth J.D.F., Joseph J.S.,
 RA Jeyaseelan K., Kint R.M., Gadre S.V., Sapatnekar S.M., Stocker M.,
 RA Strong P.N.;
 RT Tamapin, a venom peptide from the Indian red scorpion (Mesobuthus
 tamulus) that targets small conductance Ca2+-activated K+ channels
 and afterhyperpolarization currents in central neurons.";
 RL J. Biol. Chem. 277:46101-46109(2002).
 CC -1- FUNCTION: Blocks small conductance calcium-activated potassium
 channels.
 CC -1- SUBCELLULAR LOCATION: Secreted.
 CC -1- TISSUE SPECIFICITY: Expressed by the venom gland.
 CC -1- MASS SPECTROMETRY: MW=3431.4; MW ERR=0.2; METHOD=Electrospray.
 CC -1- SIMILARITY: Belongs to the short scorpion toxin family. Potassium
 channel inhibitors subfamily.
 DR PROSITE: PS01138; SCOP SHORT_TOXIN; 1.
 KW Toxin; Neurotoxin; Ionic channel inhibitor;
 KW Potassium channel inhibitor; Amidation.
 KM DISULFID 3 21 BY SIMILARITY.
 FT DISULFID 8 26 BY SIMILARITY.
 FT DISULFID 12 28 BY SIMILARITY.
 FT MOD RES 31 31 AMIDATION (PROBABLE).
 SQ SEQUENCE 31 AA; 3439 MW; E4D43DC54CA415C CRC64;

Query Match Best Local Similarity 73.8%; Score 31; DB 1; Length 31;
 Pred. No. 15; Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CXCXPH 6
 DB 26 CXCXPH 31

RESULT 12
 SCT1_CENIM ID SCT1_CENIM STANDARD; PRT; 39 AA.
 AC P59847;
 DT 10-OCT-2003 (Rel. 42, Created)
 DT 10-OCT-2003 (Rel. 42, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Hongotoxin 1 (Hgtx1).
 OS Centruroides limatus (Scorpion).
 OC Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Scorpiones;
 OC Butioidea; Butidae; Centruroides.
 OX NCBI_Taxid=244936;
 RN [1]
 RP SEQUENCE, MUTAGENESIS, AND PHARMACOLOGICAL CHARACTERIZATION.
 RC TISSUE=Venom;
 RX MEDLINE=98112806; PubMed=9446567;
 RA Koschak A., Buglianesi R.M., Milterdorfer J., Kaczorowski G.J.,
 RA Garcia M.L., Knaus H.-G.;
 RT "Subunit composition of brain voltage-gated potassium channels
 determined by hongotoxin-1, a novel peptide derived from Centruroides
 limatus venom.";
 RL J. Biol. Chem. 273:2639-2644(1998).
 RN [2]
 RP STRUCTURE BY NMR, MASS SPECTROMETRY, AND MUTAGENESIS.

```

RC TISSUE=Venom; PubMed=12009929;
RX MEDLINE=22005852;
RA Prael B., Koschak A., Trieb M., Obermaier G., Kaufmann W.A.,
RA Gerster U., Blanc E., Hahn C., Prinz H., Schuetz G., Darion H.,
RA Gruber H.J., Knans H.-G.,
RT "Synthesis, characterization, and application of cy-dye- and
RT alexa-dye-labeled homotoxin(1) analogues. The first high affinity
RT fluorescence probes for voltage-gated K+ channels.";
RL Biochem. Chem. 13:416-425(2002).
CC -1- FUNCTION: Potent selective inhibitor of Kv1.1, Kv1.2, Kv1.3
CC voltage-gated potassium channels. Weak inhibitor of Kv1.6
CC potassium channel. Does not block Kv1.4 and Kv1.5 currents.
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- TISSUE SPECIFICITY: Expressed by the venom gland.
CC -1- MASS SPECTROMETRY: MW=4219; METHOD=Electrospray.
CC -1- SIMILARITY: Belongs to the short scorpion toxin family. Potassium
CC channel inhibitor subfamily.
DR PDB: 1HLV, PRELIMINARY.
DR Prodom; PD003586; Scorpion toxins; 1.
DR PROSITE; PS01138; SCORP_SHORT_TOXIN; 1.
KM Toxin; Neurotoxin; Ionic channel inhibitor;
KM Potassium channel inhibitor; 3D-structure.
FT MUTAGEN 19 19 A->C: NO LOSS OF ACTIVITY WHEN ASSOCIATED
FT MUTAGEN 19 19 A->Y: NO LOSS OF ACTIVITY WHEN ASSOCIATED
FT MUTAGEN 37 37 Y->F: NO LOSS OF ACTIVITY WHEN ASSOCIATED
FT MUTAGEN 37 37 Y->F: NO LOSS OF ACTIVITY WHEN ASSOCIATED
SQ SEQUENCE 39 AA; 4226 MW; B4ABC83F8E7E637 CRC64;

Query Match 73.8%; Score 31; DB 1; Length 39;
Best Local Similarity 66.7%; Pred. No. 19;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 CXCPH 6
DB 34 CKCPH 39

RESULT 13
SCRM_CENMA STANDARD; PRT; 39 AA.
AC P40755;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Margatoxin (MGTX).
OS Centruroides margaritatus (Scorpion).
OC Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Scorpiones;
OC Butiidae; Butiidae; Centruroides.
NCBI_TaxID=29018;
RN [1]
RP SEQUENCE.
RC TISSUE=Venom;
RX MEDLINE=9336802; PubMed=8360176;
RA Garcia-Calvo M., Leonard R.J., Novick J., Stevens S.P.,
RA Schmalhofer W., Kaczorowski G.J., Garcia M.L.;
RT "Purification, characterization, and bioynthesis of margatoxin, a
RT component of Centruroides margaritatus venom that selectively
RT inhibits voltage-dependent potassium channels.";
RL J. Biol. Chem. 268:1886-18874(1993).
RN [2]
RP SYNTHESIS, AND DISULFIDE BONDS.
RX MEDLINE=94128107; PubMed=8297371;
RA Bednarek M.A., Bugianesi R.M., Leonard R.J., Felix J.P.,
RT "Chemical synthesis and structure-function studies of margatoxin, a
RT potent inhibitor of voltage-dependent potassium channel in human T
RT lymphocytes.";
RL Biochem. Biophys. Res. Commun. 198:619-625(1994).
RN [3]
RP STRUCTURE BY NMR.
RX MEDLINE=95092763; PubMed=7999764;
RA Johnson B.A., Stevens S.P., Williamson J.M.;
RT "Determination of the three-dimensional structure of margatoxin by

```

```

RT 1H, 13C, 15N triple-resonance nuclear magnetic resonance
RT spectroscopy.";
RL Biochemistry 33:15061-15070(1994).
CC -1- FUNCTION: Potent selective inhibitor of voltage-dependent
CC potassium channels such as Kv1.3.
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- TISSUE SPECIFICITY: Expressed by the venom gland.
CC -1- SIMILARITY: Belongs to the short scorpion toxin family. Potassium
CC channel inhibitor subfamily.
DR PIR: A48523; A48523.
DR PDB: 1MTX; 14-NOV-95.
DR InterPro; IPR001947; Scorpion toxins.
DR Pfam; PF00451; Toxin_2; 1.
DR Prodom; PD003586; Scorpion toxins; 1.
DR PROSITE; PS01138; SCORP_SHORT_TOXIN; 1.
KM Toxin; Neurotoxin; Ionic channel inhibitor;
KM Potassium channel inhibitor; 3D-structure.
FT DISULFID 7 29
FT DISULFID 13 34
FT DISULFID 17 36
FT SITE 37 39 INTERACTION WITH KV 1.3 CHANNELS
FT STRAND 6 6 (POTENTIAL).
FT HELIX 10 13
FT HELIX 14 20
FT TURN 21 21
FT TURN 23 24
FT TURN 26 30
FT TURN 31 32
FT STRAND 33 37
SQ SEQUENCE 39 AA; 4185 MW; 0FB748318014BE0D CRC64;

Query Match 73.8%; Score 31; DB 1; Length 39;
Best Local Similarity 66.7%; Pred. No. 19;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 CXCPH 6
DB 34 CKCPH 39

RESULT 14
AMCY_METEX STANDARD; PRT; 119 AA.
AC P04172;
DT 20-MAR-1987 (Rel. 04, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Amicyanin-alpha precursor.
OS Methylobacterium extorquens.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Methylobacteriaceae; Methylobacterium.
NCBI_TaxID=408;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=AM1 / NCIMB 9133;
RX MEDLINE=91358385; PubMed=1653226;
RA Chistoserdov A.Y., Tsygankov Y.D., Lidstrom M.E.;
RT "Genetic organization of methylamine utilization genes from
RT Methylobacterium extorquens AM1.";
RL J. Bacteriol. 173:5901-5908(1991).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=AM1 / NCIMB 9133;
RX MEDLINE=94292425; PubMed=8021187;
RA Chistoserdov A.Y., Chistoserdova L.V., McIntire W.S., Lidstrom M.E.;
RT "Genetic organization of the man gene cluster in Methylobacterium
RT extorquens AM1: complete nucleotide sequence and generation and
RT characteristics of man mutants.";
RL J. Bacteriol. 176:4052-4065(1994).
RN [3]
RP SEQUENCE OF 21-119.

```

RC STRAIN=AMI / NCIMB 9133;
 RX MEDLINE=86130354; PubMed=4091802;
 RA Ambler R.P., Tobari J.;
 RT "The primary structures of Pseudomonas AM1 amicyanin and
 pseudocyanin. Two new sequence classes of blue copper proteins.";
 RL Biochem. J. 232:451-457(1985).
 CC -1- FUNCTION: Primary acceptor of electrons from methyamine
 dehydrogenase. Passes those electrons on either a soluble
 cytochrome c or to pseudocyanin.
 CC -1- COFACTOR: Binds 1 copper ion per molecule.
 CC -1- PATHWAY: Methyamine utilization.
 CC -1- SUBCELLULAR LOCATION: Periplasmic.
 CC -1- SIMILARITY: Contains 1 plastocyanin-like domain.

 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).

 CC EMBL; M57963; AAA6895.1; -
 DR EMBL; L26406; AAA6937.1; -
 DR PIR; A56621; CUPSAM.
 DR HSSP; P23364; IAAC.
 DR InterPro; IPR000923; BlueCu_1.
 DR InterPro; IPR001235; Copper_blue.
 DR InterPro; IPR008972; Cupredoxin.
 DR Pfam; PF00127; copper-bind; 1.
 DR PRINTS; PR00156; COPPERBLUE.
 DR ProDom; PD00135; COPPER_BLUE.
 DR PROSITE; PS00196; COPPER_BLUE; 1.
 DR PROSITE; PS00196; COPPER_BLUE; 1.
 KW Copper; Electron transport; Periplasmic; Signal.
 FT CHAIN 1 20
 FT DOMAIN 21 119 AMICYANIN-ALPHA.
 FT METL 67 67 PLASTOCYANIN-LIKE.
 FT METL 106 106 COPPER (BY SIMILARITY).
 FT METL 109 109 COPPER (BY SIMILARITY).
 FT METL 112 112 COPPER (BY SIMILARITY).
 SQ SEQUENCE 119 AA; 12609 MW; 732PDECA8239D857 CRC64;

 Qy Query Match 73.8%; Score 31; DB 1; Length 119;
 Best Local Similarity 80.0%; Pred. No. 52;
 Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 Db 3 CXPHP 7
 106 CTRHP 110

 RESULT 15
 AMCY PARDE
 ID AMCY PARDE STANDARD; PRT; 131 AA.
 AC P23364;
 DT 01-AUG-1991 (Rel. 19, Created)
 DT 01-AUG-1991 (Rel. 19, Last sequence update)
 DT 15-MAR-2004 (Rel. 43, Last annotation update)
 DE Amicyanin precursor.
 GN MAUC OR AMI.
 OS Paracoccus denitrificans.
 OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhodobacterales;
 OC Rhodobacteraceae; Paracoccus.
 OC NCBI_TaxID=266;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=NCIMB 8944;
 RX MEDLINE=91085564; PubMed=2261991;
 RA van Spanning R.J.M., Wansell C.W., Reijnders W.N.M., Oltmann L.F.,
 RA Stouthamer A.M.;
 RT "Mutagenesis of the gene encoding amicyanin of Paracoccus
 denitrificans and the resultant effect on methyamine oxidation.";

RL FEBS Lett. 275:217-220 (1990).
 RN [2]
 RX SEQUENCE OF 27-36.
 RX MEDLINE=86243362; PubMed=3718960;
 RA Husain M., Davidson V.L.;
 RT "Properties of Paracoccus denitrificans amicyanin.";
 RL Biochemistry 25:2431-2436 (1986).
 RN [3]
 RP X-RAY CRYSTALLOGRAPHY OF COMPLEX WITH MADH.
 RX MEDLINE=92287919; PubMed=1599920;
 RA Chen L., Durlley R., Poliks B.J., Hamada K., Chen Z., Mathews F.S.;
 RA Davidson V.L., Satow Y., Huizinga E.G., Vellieux F.M.D., Hol W.G.J.;
 RT "Crystal structure of an electron-transfer complex between
 methyamine dehydrogenase and amicyanin.";
 RL Biochemistry 31:4959-4964 (1992).
 RN [4]
 RP X-RAY CRYSTALLOGRAPHY (2.4 ANGSTROMS).
 RX MEDLINE=94188715; PubMed=8140419;
 RA Chen L., Durlley R., Mathews F.S.;
 RT "Structure of an electron transfer complex: methyamine
 dehydrogenase, amicyanin, and cytochrome c551.";
 RL Science 264:86-90 (1994).
 RN [5]
 RP X-RAY CRYSTALLOGRAPHY (1.31 ANGSTROMS).
 RA Cunane L.M., Chen Z.-W., Durlley R.C.E., Mathews F.S.;
 RT "X-ray structure of the cupredoxin amicyanin, from Paracoccus
 denitrificans, refined at 1.31-A resolution.";
 RL Acta Crystallogr. D 52:676-686 (1996).
 RN [6]
 RP X-RAY CRYSTALLOGRAPHY (1.3 ANGSTROMS).
 RX MEDLINE=99080123; PubMed=9960825;
 RA Zhu Z., Cunane L.M., Chen Z.-W., Durlley R.C.E., Mathews F.S.,
 RA Davidson V.L.;
 RT "Molecular basis for interprotein complex-dependent effects on the
 redox properties of amicyanin.";
 RL Biochemistry 37:17128-17136 (1998).
 CC -1- FUNCTION: Primary acceptor of electrons from methyamine
 dehydrogenase. Passes those electrons on either a soluble
 cytochrome c or to pseudocyanin.
 CC -1- COFACTOR: Binds 1 copper ion per molecule.
 CC -1- PATHWAY: Methyamine utilization.
 CC -1- SUBCELLULAR LOCATION: Periplasmic.
 CC -1- SIMILARITY: Contains 1 plastocyanin-like domain.

 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).

 CC EMBL; X55665; CAA39199.1; -
 DR EMBL; A24407; A24407.
 DR PIR; S12972; S12972.
 DR PDB; 1AAJ; 31-OCT-93.
 DR PDB; 1AAN; 31-OCT-93.
 DR PDB; 1MDA; 31-OCT-93.
 DR PDB; 2MTA; 31-JAN-94.
 DR PDB; 1AAC; 08-MAR-96.
 DR PDB; 1BXA; 07-OCT-98.
 DR PDB; 2RAC; 07-OCT-98.
 DR PDB; 1MG2; 11-DEC-02.
 DR InterPro; IPR000923; BlueCu_1.
 DR InterPro; IPR001235; Copper_blue.
 DR InterPro; IPR008972; Cupredoxin.
 DR Pfam; PF00127; copper-bind; 1.
 DR PRINTS; PR00156; COPPERBLUE.
 DR ProDom; PD00135; COPPER_BLUE; 1.
 DR PROSITE; PS00196; COPPER_BLUE; 1.
 KW Copper; Electron transport; Periplasmic; Signal; 3D-structure.
 FT SIGNAL 1 26
 FT CHAIN 27 131 AMICYANIN.

FT	DOMAIN	27	131	PLASTOCYANIN-LIKE.
FT	METAL	79	79	COPPER (BY SIMILARITY).
FT	METAL	118	118	COPPER (BY SIMILARITY).
FT	METAL	121	121	COPPER (BY SIMILARITY).
FT	METAL	124	124	COPPER (BY SIMILARITY).
FT	STRAND	29	30	
FT	STRAND	37	38	
FT	HELIX	39	41	
FT	TURN	44	45	
FT	STRAND	47	52	
FT	TURN	53	54	
FT	STRAND	55	56	
FT	STRAND	60	63	
FT	TURN	65	66	
FT	STRAND	68	73	
FT	STRAND	79	79	
FT	STRAND	82	83	
FT	TURN	85	87	
FT	STRAND	93	94	
FT	STRAND	98	98	
FT	TURN	100	101	
FT	STRAND	102	108	
FT	STRAND	112	118	
FT	TURN	119	120	
FT	STRAND	121	131	
SQ	SEQUENCE	131	AA; 13983 MM; F7352A865FD089DA CRC64;	

Query Match 73.8%; Score 31; DB 1; Length 131;
 Best Local Similarity 80.0%; Pred. No. 57;
 Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 Qy 3 CXPMP 7
 Db 118 CTPMP 122

Search completed: April 8, 2004, 11:11:04
 Job time : 5.83333 secs